BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

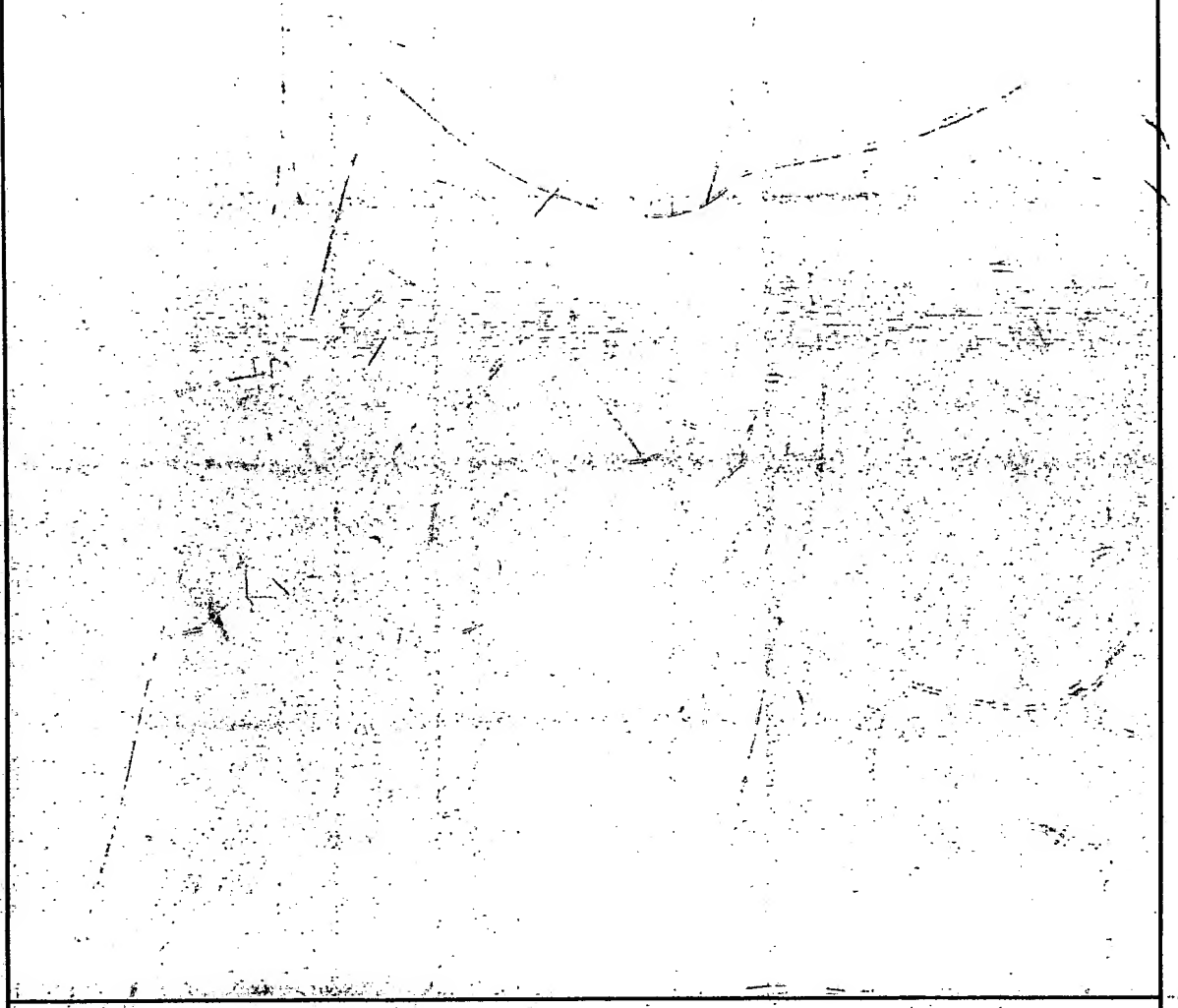
SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____

Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevent citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevent claim(s).



STAFF USE ONLY

Date completed: 06-30-03

Searcher: Beverly C4994

Terminal time: 20

Elapsed time: _____

CPU time: _____

Total time: 28

Number of Searches: _____

Number of Databases: 1

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

Vendors

_____ IG

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

✓ DARC/Questel

_____ Other CGN

PTO-1590 (9-90)

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:50:54 ; Search time 1211.48 Seconds
(without alignments)
5454.265 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIKM.....AYLTKXXNDLREQGQPSQRT 408

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09208629/runat_24062003_101625_12930/app_query.fasta_1.1166
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629_@CGN_1_1_2013_@runat_24062003_101625_12930 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	947.5	43.9	581	13	BM258063	BM258063 522739 MA
	2	790.5	36.6	554	13	BM255484	BM255484 517232 MA
	3	724	33.6	585	17	AZ820748	AZ820748 2M0093H12
	4	707	32.8	529	9	AA177828	AA177828 mt07f02.r
	5	615	28.5	495	9	AL713459	AL713459 DKFZp686N
	6	563.5	26.1	839	14	BQ897419	BQ897419 AGENCOURT
	7	491.5	22.8	685	9	AL660446	AL660446 AL660446
	8	476.5	22.1	938	17	CNS04SUP	AL305674 Tetraodon
	9	464.5	21.5	765	13	BI183645	BI183645 UNL-P-FN-
	10	453.5	21.0	821	14	BM946931	BM946931 UI-M-EH0p
	11	437	20.3	801	13	BG924078	BG924078 602823635
	12	436.5	20.2	624	13	BI394593	BI394593 pgp1n.pk0
	13	423	19.6	860	9	AJ446632	AJ446632 AJ446632
	14	422	19.6	800	9	AJ452624	AJ452624 AJ452624
	15	421.5	19.5	808	9	AJ456719	AJ456719 AJ456719
	16	418.5	19.4	890	9	AJ456561	AJ456561 AJ456561
	17	416.5	19.3	741	9	AJ453402	AJ453402 AJ453402
c	18	413	19.1	265	10	BE146323	BE146323 MR0-HT020
	19	412.5	19.1	845	9	AJ456135	AJ456135 AJ456135
c	20	409.5	19.0	602	13	BM439733	BM439733 pgr1n.pk0
	21	409.5	19.0	747	9	AJ446611	AJ446611 AJ446611
	22	409.5	19.0	773	9	AJ451396	AJ451396 AJ451396
	23	408.5	18.9	749	9	AJ447002	AJ447002 AJ447002
	24	405.5	18.8	766	9	AJ450835	AJ450835 AJ450835
	25	401.5	18.6	740	9	AJ450921	AJ450921 AJ450921
	26	401.5	18.6	760	9	AJ456491	AJ456491 AJ456491
	27	400.5	18.6	730	9	AJ455912	AJ455912 AJ455912
c	28	391.5	18.2	817	17	CNS03DLV	AL239260 Tetraodon
	29	389.5	18.1	2020	11	BC013202	BC013202 Homo sapi
	30	383.5	17.8	746	9	AJ450259	AJ450259 AJ450259
	31	383.5	17.8	798	9	AJ451155	AJ451155 AJ451155
	32	377.5	17.5	724	9	AJ447835	AJ447835 AJ447835
	33	376	17.4	633	13	BJ068617	BJ068617 BJ068617
	34	373.5	17.3	591	13	BI984999	BI984999 fu13d01.y
	35	372.5	17.3	682	9	AJ453748	AJ453748 AJ453748
	36	370.5	17.2	668	10	BE005953	BE005953 RC0-BN012
	37	369.5	17.1	531	12	BF426487	BF426487 df70d03.y
	38	368.5	17.1	1199	11	AF345566	AF345566 Homo sapi
	39	364	16.9	904	9	AL532537	AL532537 AL532537
	40	364	16.9	931	9	AL547762	AL547762 AL547762
	41	364	16.9	1766	11	AK008952	AK008952 Mus muscu
	42	363	16.8	2542	11	AK017378	AK017378 Mus muscu
	43	363	16.8	3001	11	AK005013	AK005013 Mus muscu
	44	362.5	16.8	746	9	AJ456692	AJ456692 AJ456692
c	45	361.5	16.8	542	9	AI942392	AI942392 wo77g12.x

ALIGNMENTS

RESULT 1
BM258063
LOCUS BM258063 581 bp mRNA linear EST 17-DEC-2001
DEFINITION 522739 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BM258063
VERSION BM258063.1 GI:17893662
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 581)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett

```
          ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
          Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
          Keele,J.W.
  TITLE   Sequence evaluation of four pooled-tissue normalized bovine cDNA
          libraries and construction of a gene index for cattle
  JOURNAL Genome Res. 11 (4), 626-630 (2001)
  MEDLINE 21180013
COMMENT   Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@email.marc.usda.gov
          Single pass sequencing. Bases called and alt_trimmed with phred
          v0.980904.e. Vector identified by cross_match with the -minscore 18
          and -minmatch 12 options.
          PCR PRimers
          FORWARD: AGGAAACAGCTATGACCAT
          BACKWARD: GTTTTCCCAGTCACGACG
          Plate: 126  row: D  column: 8
          Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 3BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from marrow, alveolar
                     macrophage, ovary, fetal semitendonosus muscle, and fetal
                     longissimus muscle."
BASE COUNT      122 a    171 c    117 g    171 t
ORIGIN

Alignment Scores:
Pred. No.:              6.11e-91        Length:         581
Score:                  947.50          Matches:        176
Percent Similarity:     95.88%          Conservative:   10
Best Local Similarity:  90.72%          Mismatches:     7
Query Match:            43.93%          Indels:         1
DB:                     13              Gaps:           1

US-09-208-629F-6 (1-408) x BM258063 (1-581)

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
              ||||||||||||:::|||   |||||||||||||||||||||||||||:::||||||
Db          2 ATGGGGTACCTGAGCAGCCCTTTAAGTACCAAATTGATACCCGCCATCTACATCCTGGTG 61

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArg 143
              |||   |||||:::||||||||||||||||||||||||||   ||||||||||||
Db         62 TTTGCAGTAGGTATGCCGGCCAATGCGGTGACCCTGTGGATGCTC---TTCAGGACCAGA 118

Qy        144 SerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
              :::|||     |||:::||||||||||||||||||||||||||||||||||||||||
Db        119 ACCATCCGTATGACCATCTTCTACACCAACCTGGCCATTGCAGACTTTCTGTTTTGTGTT 178

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        179 ACACTGCCCTTTAGAATAGCTTACCATCTCAATGGGAACAACTGGGTATTTGGCGAGGTC 238

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 ATGTGCCGGGCCACCACGGTCATCTTCTATGGCAACATGTATTGCTCCATTCTGCTCCTC 298

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GCCTGCATCAGTATCAACCGCTACTTAGCCATTGTTCATCCTTTCACTTACCGGGGACTG 358

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
              ||||||   ||||||||||:::|||||||||||||||||   ||||||||||||||||
Db        359 CCCAAGCGCACCTACGCCTTGCTCACATGTGGACTGGTGTGGACAACGGTTTTCTTATAC 418

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db        419 ATGCTGCCGTTTTTCATTCTGAAGCAGGAGTACTATCTTGTCCAGCAAGACATTACCACC 478

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TGCCACGATGTCCACAACACATGCGAGTCCTCGTCGCCCTTCCAGCTCTACTACTTCATC 538

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIle 297
              |||||||||||||||||||||||||||||||:::::|||
Db        539 TCCTTGGCATTCTTTGGATTCTTAATCCCATTTCTGGTCATT 580

RESULT 2
BM255484
LOCUS       BM255484                 554 bp    mRNA    linear   EST 17-DEC-2001
DEFINITION  517232 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BM255484
VERSION     BM255484.1  GI:17891083
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 117  row: K  column: 11
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .554
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 3BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from marrow, alveolar
                     macrophage, ovary, fetal semitendonosus muscle, and fetal
                     longissimus muscle."
BASE COUNT      140 a    148 c    118 g    148 t
ORIGIN

Alignment Scores:
Pred. No.:              3.41e-74        Length:         554
Score:                  790.50          Matches:        151
Percent Similarity:     89.13%          Conservative:   13
Best Local Similarity:  82.07%          Mismatches:     19
Query Match:            36.65%          Indels:         1
DB:                     13              Gaps:           1

US-09-208-629F-6 (1-408) x BM255484 (1-554)

Qy         17 ValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr 36
```

```
Db          6 CTGATCAAGATGAGAGCCGTCATCTTTGCAGCCATTGGAGCACTACTTCTGTCGCCTGCT 65

Qy         37 PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIle 56

Db         66 TCCTGTCAAAGCGGCATGGAATATGATGCAGACAACTTGGCAAAGCCAACCTTGTCTATT 125

Qy         57 LysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGlu 76

Db        126 AAGACCTTCCGTGGAGCTCCCCAAAATTCTTTTGAAGAGTTCCCCCTTTCTGCCATAGAA 185

Qy         77 GlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHis 96

Db        186 GGCTGGACAGGAACCACCAAAACTGTAAAAATCAAATGCCCTGAAGAACTTGATTCAAAT 245

Qy         97 LeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIle 116

Db        246 CTCCATGTGAATAATGCTACCATGGGGTACCTGAGCAGCCCTTTAAGTACCAAATTGATA 305

Qy        117 ProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrp 136

Db        306 CCCGCCATCTACATCCTGGTGTTTGCAGTAGGTATGCCGGCCAATGCGGTGACCCTGTGG 365

Qy        137 MetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIle 156

Db        366 ATGCTC---TTCAGGACCAGAACCATCCGTATGACCATCTTCTACACCAACCTGGCCATT 422

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176

Db        423 GCAGACTTTCTGTTTTGTGTTACACTGCCCTTTAGAATAGCTTACCATCTCAATGGGAAC 482

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196

Db        483 AACTGGGTATTTGGCGAGGTCATGTGCCGGGCCACCACGGTCATCTTCTATGGCAACATG 542

Qy        197 TyrCysSerIle 200

Db        543 TATTGCTCCATT 554
```

RESULT 3
AZ820748
LOCUS AZ820748 585 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0093H12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0093H12 F, DNA sequence.
ACCESSION AZ820748
VERSION AZ820748.1 GI:12990656
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 585)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0093 row: H column: 12
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 585.
FEATURES Location/Qualifiers

source 1. .585
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0093H12"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 153 a 168 c 117 g 147 t
ORIGIN

Alignment Scores:

Pred. No.:	4.7e-67	Length:	585
Score:	724.00	Matches:	134
Percent Similarity:	82.90%	Conservative:	26
Best Local Similarity:	69.43%	Mismatches:	31
Query Match:	33.57%	Indels:	2
DB:	17	Gaps:	0

US-09-208-629F-6 (1-408) x AZ820748 (1-585)

```
Qy         44 AsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaPro 63

Db         11 AATGTTTCAGACAACTCAGCAAAGCCAACCTTAACTATTAAGAGTTTTAATGGGGGTCCC 70

Qy         64 ProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIle 83

Db         71 CAAAATACCTTTGAAGAATTCCCACTTTCTGACATAGAGGGCTGGACAGGAGCCACCACA 130

Qy         84 ThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThr 103

Db        131 ACTATAAAAGCGGAGTGTCCCGAGGACAGTATTTCAACTCTCCACGTGAATAATGCTACC 190

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123

Db        191 ATAGGATACCTGAGAAGTTCCTTAAGTACCCAAGTGATACCTGCCATCTACATCCTGCTG 250

Qy        124 PheVal-ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrAr 143

Db        251 TTTGTGAGTTGGTGTACCAGCCAACATCGTGACCCTGTGGAAACTCTCCTTAAGGACCAA 310

Qy        143 gSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVa 163

Db        311 ATCCATCAGTCTGGTCATCTTTCACACCAACCTGGCCATCGCAGATCTCCTTTTCTGTGT 370

Qy        163 lThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe-GlyGluV 183

Db        371 CACACTGCCATTTAAGATCGCCTACCATCTCAATGGCAACAACTGGGTATTTGGGCGAGG 430

Qy        183 alLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuL 203

Db        431 TCACGTGCCGGATCACCACGGTTGTTTTCTACGGCAACATGTACTGCGCTATCCTGATCC 490

Qy        203 euAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyL 223
```

```
Db        491 TCACCTGCATGGGCATCAACCGCTACCTGGCCACGGCTCACCCTTTCACATACCAGAAGC 550

Qy        223 euProLysHisThrTyrAlaLeuValThrCysGly 234
Db        551 TGCCCAAACGCAGCTTCTCCATGCTCATGTGTGGC 585

RESULT 4
AA177828
LOCUS       AA177828                529 bp    mRNA    linear   EST 16-FEB-1997
DEFINITION  mt07f02.r1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:620379
            5' similar to TR:G663021 G663021 PROTEINASE ACTIVATED RECEPTOR 2.
            ;, mRNA sequence.
ACCESSION   AA177828
VERSION     AA177828.1  GI:1759090
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:381203
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 494.
FEATURES             Location/Qualifiers
     source          1. .529
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:620379"
                     /clone_lib="Soares mouse 3NbMS"
                     /sex="male"
                     /tissue_type="Spleen"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through
                     three rounds of normalization, and was constructed by
                     Bento Soares and M.Fatima Bonaldo."
BASE COUNT      110 a    168 c    107 g    144 t
ORIGIN

Alignment Scores:
Pred. No.:              2.67e-65        Length:         529
Score:                  707.00          Matches:        125
Percent Similarity:     84.57%          Conservative:   23
Best Local Similarity:  71.43%          Mismatches:     27
Query Match:            32.78%          Indels:         0
DB:                     9               Gaps:           0

US-09-208-629F-6 (1-408) x AA177828 (1-529)

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
Db          3 AGTTCCTTAAGTACCCAAGTGATACCTGCCATCTACATCCTGCTGTTTGTGGTTGGTGTA 62

Qy        129 ProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThr 148
Db         63 CCAGCCAACATCGTGACCCTGTGGAAACTCTCCTTAAGGACCAAATCCATCAGTCTGGTC 122

Qy        149 ValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLys 168
Db        123 ATCTTTCACACCAACCTGGCCATCGCAGATCTCCTTTTCTGTGTCACACTGCCATTTAAG 182

Qy        169 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThr 188
Db        183 ATCGCCTACCATCTCAATGGCAACAACTGGGTATTTGGCGAGGTCACGTGCCGGATCACC 242

Qy        189 ThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIle 208
Db        243 ACGGTTGTTTTCTACGGCAACATGTACTGCGCTATCCTGATCCTCACCTGCATGGGCATC 302

Qy        209 AsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyr 228
Db        303 AACCGCTACCTGGCCACGGCTCACCCTTTCACATACCAGAAGCTGCCCAAACGCAGCTTC 362

Qy        229 AlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePhe 248
Db        363 TCCATGCTCATGTGTGGCATGGTGTGGGTCATGGTTTTCTTATACATGCTGCCCTTTGTC 422

Qy        249 IleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHis 268
Db        423 ATCCTGAAGCAGGAGTACCACCTCGTCCACTCCGAGATCACCACCTGCCACGATGTCGTT 482

Qy        269 AsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
Db        483 GACGCGTGCGAGTCCCCATCATCCTTCCGATTCTACTACTTCGTC 527

RESULT 5
AL713459
LOCUS       AL713459                495 bp    mRNA    linear   EST 22-MAR-2002
DEFINITION  DKFZp686N1799_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686N1799 5', mRNA sequence.
ACCESSION   AL713459
VERSION     AL713459.1  GI:19696815
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Wambutt R
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686N1799) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .495
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686N1799"
                     /clone_lib="686 (synonym: hlcc3)"
                     /tissue_type="human skeletal muscle"
                     /dev_stage="adult"
```

```
                     /lab_host="DH10B"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
BASE COUNT      132 a    128 c    107 g    128 t
ORIGIN

Alignment Scores:
Pred. No.:                1.65e-55       Length:          495
Score:                    615.00         Matches:         121
Percent Similarity:       99.18%         Conservative:    0
Best Local Similarity:    99.18%         Mismatches:      1
Query Match:              28.51%         Indels:          0
DB:                       9              Gaps:            0

US-09-208-629F-6 (1-408) x AL713459 (1-495)

Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        129 TGCTCCATGATTTTACAGATTTCATAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 188

Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 248

Qy         41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 308

Qy         61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 368

Qy         81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 428

Qy        101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 488

Qy        121 LeuLeu 122
              ||||||
Db        489 CTCCTG 494

RESULT 6
BQ897419
LOCUS       BQ897419                 839 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8671134 NICHD XGC Emb4 Xenopus laevis cDNA clone
            IMAGE:4681483 5', mRNA sequence.
ACCESSION   BQ897419
VERSION     BQ897419.1  GI:22289433
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 839)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Igor Dawid
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Plate: LLAM10561  row: k  column: 20
            High quality sequence stop: 737.

FEATURES             Location/Qualifiers
     source          1. .839
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:4681483"
                     /clone_lib="NICHD XGC Emb4"
                     /dev_stage="embryo, stage 31-32"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: whole embryo; Vector: pCMV-SPORT6; Site_1:
                     NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
                     dT. Average insert size 2.1 kb. Constructed by Life
                     Technologies. Note: This is a Xenopus Gene Collection (XGC
                     ) library."
BASE COUNT      217 a    209 c    180 g    232 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:                1.1e-49        Length:          839
Score:                    563.50         Matches:         111
Percent Similarity:       62.08%         Conservative:    38
Best Local Similarity:    46.25%         Mismatches:      74
Query Match:              26.12%         Indels:          17
DB:                       14             Gaps:            4

US-09-208-629F-6 (1-408) x BQ897419 (1-839)

Qy         30 LeuLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAsp---------- 45
              |||:::|||||||||  :::          |||:::::::::|||
Db        122 CTTGTCCTACTCCTTTTGTCCGGGGCTGCTCTAGGGCTCCAAGATGACGATATGCTGGAG 181

Qy         46 ThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProPro--- 64
                |||     |||  |||           ||||||||||||||
Db        182 AAAAATGAATCTGCTGTACCC----------AAGACATTTCGGGGAAAGAAAGATGGT 229

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThr------ 82
                 :::|||:::         |||   :::  |||           |||
Db        230 GGGGACTATGAACAGCTTTCCCATAGCATTATCCACGGCCTTCCAGAGACCACTACCAAC 289

Qy         83 ----------------IleThrValLysIleLysCysProGluGluSerAlaSerHis 96
                           |||:::                 :::  :::  :::
Db        290 AACCATCTTCTTCTAAGAATCTCCAGCAATAATGCCTTTAAAGATGCCAACAAAAACACC 349

Qy         97 LeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIle 116
              :::  |||  |||  |||:::||||||||:::     :::|||      ||||||
Db        350 ATCAAAGTCAGCAACTTGACCTTGGGGTACCTGAGTGGCAAAGTAAGCAGGGCACTGATC 409

Qy        117 ProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrp 136
              |||  ||||||::::::|||  ::::::||||||||||||||||||:::  ||||||
Db        410 CCAGGCATTTATATCATTGTTGTGCTCATTGGAGTACCAGCTAATGCCATTGTGTTGTGG 469

Qy        137 MetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIle 156
              ||||||||  :::   |||:::||||||||||||  |||  :::||||||
Db        470 ATGTTGTTCCATCAGGTGACGTCTGTGTGCACAACTGTGCTCTACGCCAGCTTAGCAACC 529

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
              :::|||  |||:::|||:::  ||||||||||||||  |||||||  ||||||
Db        530 TCTGACTTGCTCTATTGCCTCATGCTGCCCTTCAAGATAGTTTATCACTTAACTGGCAAC 589

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
              ||||||:::||||||||  :::||||||      |||:::      |||  ||||||
Db        590 AACTGGATTTTTGGAGAAACCATGTGCCGGACTATGACAATCTTCTTGTACTTTAACATG 649

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
              ||||||||||||||||||  |||  ||||||:::||||||:::||||||||||||
Db        650 TACTGCTCTATTTTGCTCCTGATGTGCTTTAGCATCAGCCGCTATGTTGCCATTGTCCAT 709

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuVal 236
              ||||||  ||||||  ||||||      ||||||  ::::::  |||  ||||||
Db        710 CCATTTATATACAGGAGCTTGCCTANAGGGACCTATACGGTGCTTCTGTGCCTGCTTGTG 769

Qy        237 TrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeu 256
                     :::||||||:::||||||||||||  :::|||  |||:::|||
Db        770 TGTACCATTGCACTGGTCTACATGATCCCATTTTTCATAACCAGGCAAACATATCACTTG 829
```

```
RESULT 7
AL660446
LOCUS       AL660446                685 bp    mRNA    linear   EST 13-DEC-2001
DEFINITION  AL660446 XGC-neurula Silurana tropicalis cDNA clone TNeu043b15 5',
            mRNA sequence.
ACCESSION   AL660446
VERSION     AL660446.1  GI:17674619
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (10_2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Huckle E
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TNeu043b15.sp6
            Sequencing primer: SP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="TNeu043b15"
                     /clone_lib="XGC-neurula"
                     /dev_stage="neurula"
                     /lab_host="Escherichia coli DH10B"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from neurula.
                     EcoRI-NotI cut cDNA was then ligated into pCS107 with
                     EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT      177 a    169 c    152 g    184 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:              4.09e-42        Length:         685
Score:                  491.50          Matches:        103
Percent Similarity:     61.09%          Conservative:   32
Best Local Similarity:  46.61%          Mismatches:     75
Query Match:            22.79%          Indels:         11
DB:                     9               Gaps:           5

US-09-208-629F-6 (1-408) x AL660446 (1-685)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||         :::   |||   |||     :::       :::|||   |||
Db         26 GGNANAAGTACCACTGAATGGAAAAGGCTGAAGATGAGTCCCTTCCTTGTCCTATTCCTT 85

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
              |||                  :::   ||||||           :::
Db         86 TTGTGTGGGGTTGCTCTGGGGCTCAAAGATGAGGATACGCTGGAGAAAAATGAATCTGCT 145

Qy         54 LeuProIleLysThrPheArgGlyAlaProPro---AsnSerPheGluGluPheProPhe 72
              :::|||   |||||||||||||||                 :::|||||
Db        146 GTCCCC---AAGACATTTCGGGGAAAGAAAGAAGGTGGGGAATATGAAGAGCTTTCCCAT 202

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThrIleThr----------ValLysIleLys 88
              |||   |||   |||          |||                 :::::::|||
Db        203 AGCATTCTGCATGGCCTGCCAGACACCACAACCAATAATAATCCTTTTCTAAGAATCTCC 262

Qy         89 CysPro--------GluGluSerAlaSerHis------LeuHisValLysAsnAlaThr 103
                          :::   ::::::   |||           |||   |||:::|||
Db        263 AGCAATAATGCCTTTAAAGTCGCCAGCGAACACACCATCACTAAAGTCAGCAATTCAACC 322

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
              :::   ||||||:::       :::|||   :::|||||||||   ||||||:::::|||
Db        323 TTGGTGTACCTAAGTGGCAAAGTAAGCAGGGAACTGATACCGGGCATCTATATCATTGTT 382

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArg 143
                 ::::::|||||||||||||||||||:::   |||||||||||||||   :::
Db        383 GTGCTCATTGGTGTGCCAGCCAATGCCTTGGTACTATGGATGCTGTTCCATCAGGTGACA 442

Qy        144 SerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
              |||:::|||||||||||   |||   :::||||||   :::|||   |||:::|||:::
Db        443 TCTGTGTGCACCACTGTGCTCTACGCCAGCTTAGCCACCTCTGACTTGCTCTATTGCCTC 502

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
              ||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        503 ATGCTGCCTTTTAAGATAGCTTATCACTTAAACGGCAACAACTGGATTTTTGGAGAAACC 562

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
              :::|||||||||   |||:::      |||   |||||||||||||||||||||||||||
Db        563 ATGTGCCGGGCCATGACCATTTTCTTGTACTTTAACATGTACTGCTCCATTCTGCTCCTG 622

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
              |||   |||||||||||||||:::|||||||||||||||||||   |||||   |||
Db        623 ATGTGCTTTAGCATCAACCGCTACGTTGCCATTGTCCATCCATTTATATACAGGAGCTTG 682

Qy        224 Pro 224
              |||
Db        683 CCG 685

RESULT 8
CNS04SUP
LOCUS       CNS04SUP                938 bp    DNA     linear   GSS 24-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            001A08 of library H from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL305674
VERSION     AL305674.1  GI:8199451
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 938)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 938)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 938)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .938
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="001A08"
                     /clone_lib="H"
                     /note="Genoscope sequence ID : XC0AH001BA04XD1~end : T7"
BASE COUNT      161 a    307 c    261 g    207 t      2 others
ORIGIN
```

```
Alignment Scores:
Pred. No.:              2.56e-40        Length:         938
Score:                  476.50          Matches:        97
Percent Similarity:     58.30%          Conservative:   47
Best Local Similarity:  39.27%          Mismatches:     92
Query Match:            22.09%          Indels:         12
DB:                     17              Gaps:           4

US-09-208-629F-6 (1-408) x CNS04SUP (1-938)

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
              :::::::::|||      |||||||||:::   |||   ||||||   :::   :::|||
Db         17 TCCTCCATCTTCATGGCCAACCTGGCCGTGGC-GACCTGCTCTTCGTCATCTGGGTCCCG 75

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||||||||||||||||||:::||||||:::||||||   ||||||:::
Db         76 CTGAAAATCGCCTACCACCTCAACGGGAACGACTGGGTCTACGGCGAGTCCCTGTGCAAG 135

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                          |||||||||||||||||||||||||      :::||||||||||
Db        136 GTCCTGGTGGCCTTCTTCTACGGCAACATGTACTGCTCCATCGCCTTCATCGCGTGCATC 195

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
              |||:::   |||      |||::::::||||||          ||||||
Db        196 AGCGTCCAGCGCTGCAAGGCCGTGCTCCACCCGCTCGACAAGAGGGGG------GGCAGC 249

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
                  |||:::                :::|||      |||:::|||      :::|||
Db        250 GGTGCGGCCGTGGGCGTGTCCGCGGCCATCTGGCTGCTGGTGTGGCTCATCACCGTTCCT 309

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
                 |||:::      |||:::      :::             |||   ||||||||||||
Db        310 CTGTTCCTGTACGATCAGCAGGTCCTGGTGACAAACCTGGGCATCCGCACCTGCCACGAC 369

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAla 286
              |||      |||      |||                 ||||||:::::::::
Db        370 GTC------ACCAGGCCCAGCCAGAAGGAGACGGCTGCAGGCTACTTCCTGACCATGGGA 423

Qy        287 PhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArg 306
                  ||||||::::::|||   ::::::   |||      ||||||   :::::::::
Db        424 ACGCTGGGCTTTATTGTTCCCGCCCTCGTGTGCATCGTGTCCTACGCTCTCATGCTCAAG 483

Qy        307 ThrLeuAsn-----------------AlaTyrAspHisArgTrpLeuTrpTyrValLys 320
                 |||                                                |||
Db        484 GCCCTCAGGAAAAGCATGACGGATCCGACCATCACCCGGATGCGGCGGAAGGCCGTGGCC 543

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
                  :::::::::||||||:::|||   :::||||||   |||||||||:::::|||
Db        544 TTGATTGTBATAGTCCTGGTGATGTTCCTGGTCTGCTTCACTCCCAGCAACGTCATGCTG 603

Qy        341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
              ::::::|||::::::              :::   :::   ||||||      |||:::
Db        604 CTGGTGCATTACTCCCTCCTTTTGAGCGGCTCTCCCAACACGCTGTACGGCGCCTACGTC 663

Qy        361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
                  |||||||||   |||||||||||||||||||||||||||||:::   :::
Db        664 ACCACGCTGTGCCTGGCCAGCCTCAACAGCTGCCTGGACCCCTTCCTGTACTACTTCATC 723

Qy        381 SerLysThr---ArgAsnHis 386
              |||:::      |||   |||
Db        724 TCAGAGGACTTCAGGGAGCAC 744

RESULT 9
BI183645
LOCUS       BI183645                765 bp    mRNA    linear   EST 10-JUL-2001
DEFINITION  UNL-P-FN-cb-d-02-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
            UNL-P-FN-cb-d-02-0-UNL 3', mRNA sequence.
ACCESSION   BI183645
VERSION     BI183645.1  GI:14658054
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   Caetano,A.R., Johnson,R.K. and Pomp,D.
  TITLE     Generation and sequence characterization of a normalized cDNA
            library from swine ovarian follicles
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Pomp, D
            Department of Animal Science
            University of Nebraska, Lincoln
            Lincoln, NE 68583-0908, USA
            Tel: 402 472 6416
            Fax: 402 472 6362
            Email: dpomp@unl.edu
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message.
            Seq primer: M13 -29
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Sus scrofa"
                     /strain="University of Nebraska, Lincoln Swine Selection
                     Lines"
                     /db_xref="taxon:9823"
                     /clone="UNL-P-FN-cb-d-02-0-UNL"
                     /clone_lib="UNL-P-FN"
                     /dev_stage="ADULT"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UNL-P-FN
                     library is a normalized library representing porcine
                     ovarian follicles, ranging between 2.0 to 10.0 mm in
                     diameter, collected  during 7 days of the follicular phase
                     of the pig estrous cycle. This  library was derived from
                     the library UNL-P-F2. The tag is a string of  5-6
                     nucleotides present between the Not I site and the
                     oligo-dT track.  The library was constructed as described
                     by Bonaldo, Lennon and Soares,  Genome Research 6: 791-806
                     , 1996.
                     TAG_SEQ=None found"
BASE COUNT      142 a    217 c    186 g    219 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              3.67e-39        Length:         765
Score:                  464.50          Matches:        99
Percent Similarity:     54.90%          Conservative:   41
Best Local Similarity:  38.82%          Mismatches:     94
Query Match:            21.53%          Indels:         21
DB:                     13              Gaps:           6

US-09-208-629F-6 (1-408) x BI183645 (1-765)

Qy         95 SerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLys 114
              |||      |||        :::      ||||||      |||:::|||
Db          9 TCAGGCTTTTCCGTGGATGAATTTTCTGCTTCTGTCCTCACTGGGAAATTGACTACTGTC 68

Qy        115 LeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThr 134
                 :::|||   :::|||   :::|||||||||||||||:::|||:::|||   :::
Db         69 TTTCTTCCGGTTGTCTACACGATTGTGTTTGTGGTTGGTTTGCCAAGCAATGGCATGGCC 128

Qy        135 LeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsn 153
              ||||||:::      |||||||||:::                 |||   |||      |||
Db        129 CTGTGGGTCTTTCTTTTCCGAACGAAGAAGAAGCATCCCGCTGTGATTTACATGGCCAAT 188

Qy        154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
              ||||||:::||||||   |||      :::      |||   |||||||||||||:::
Db        189 CTGGCCTTGGCGGACCTCCTCTCTGTTATCTGGTTCCCTTTGAAGATCGCCTATCACATA 248

Qy        174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
              :::|||||||||||||:::||||||   ||||||:::                ||||||
Db        249 CATGGCAACAACTGGGTTTACGGGGAATCTCTTTGCAAGGTGCTCATTGGCTTTTTCTAC 308
```

```
Qy        194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
              ||||||||||||||||||||||||   :::   |||:::|||:::   ||||||
Db        309 GGCAACATGTACTGCTCCATCCTCTTCATGACCTGCCTCAGCGTTCAGAGGTACTGGGTC 368

Qy        214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233
              ||||||:::|||      :::          :::             |||:::
Db        369 ATTGTGAATCCCATGGTGCACCCCAGGAAGAAGGCAAACGTTGCCATCGGCGTCTCCCTG 428

Qy        234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
              |||   :::|||      :::   |||      :::|||   :::::::::||||||
Db        429 GGA---ATATGGCTGCTGATTCTGCTGGTGACCATCCCCCTGTATGTCGTGAAGCAGACT 485

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
              |||:::           ||||||||||||||||||||
Db        486 CTCTACATCCCAGCCCTTCACATCACAACCTGTCATGACGTC----------------- 527

Qy        274 SerSerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhe 287
              |||   ::::::                        ||||||:::||||||||
Db        528 ---TTGCCCGAGGAGGTGTTGGTAGGGGACATGTTCAATTATTTCCTCTCTCTGGCCATC 584

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
              ||||||   |||      |||             |||      :::||||||||||
Db        585 GGAGTCTTCCTGTTCCCCGCCTTCCTCACGGCCGCTGCCTACGTGCTGATGATCAGGACG 644

Qy        308 Leu------AsnAlaTyrAspHis-----------ArgTrpLeuTrpTyrValLysAla 321
              |||      :::|||   |||              :::            :::|||
Db        645 CTGCGGTCTTCTGCCATNGACGAAAACTCAGGAAAGAAGAGGCAGCGAGCCATTAAGCTC 704

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::   :::|||   ::::::   |||||||||   ||||||
Db        705 ATCATCACCGTCCTGGCCATGTACCTGATCTGCTTCACGCCTAGC 749

RESULT 10
BM946931
LOCUS       BM946931                821 bp    mRNA    linear   EST 14-MAR-2002
DEFINITION  UI-M-EH0p-bvn-k-05-0-UI.r1 NIH_BMAP_EH0p Mus musculus cDNA clone
            IMAGE:5694460 5', mRNA sequence.
ACCESSION   BM946931
VERSION     BM946931.1  GI:19430516
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 821)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, Univeristy of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .821
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5694460"
                     /clone_lib="NIH_BMAP_EH0p"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 18.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured mRNa was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was size selected according to mRNA size fraction,
                     ligated with EcoR I adaptor, digested with Not I, and then
                     cloned directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     , is CAGCCACGAC. This library was created for the
                     University of Iowa Mouse Brain Molecular Anatomy Project
                     (BMAP): 'Gene Discovery in the Developing Mouse Nervous
                     System', supported by National Institute of Mental Health
                     (NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT      127 a    261 c    207 g    226 t
ORIGIN

Alignment Scores:
Pred. No.:                6.07e-38        Length:          821
Score:                    453.50          Matches:         99
Percent Similarity:       57.39%          Conservative:    64
Best Local Similarity:    34.86%          Mismatches:      90
Query Match:              21.02%          Indels:          32
DB:                       14              Gaps:            10

US-09-208-629F-6 (1-408) x BM946931 (1-821)

Qy        116 IleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsn------AlaVal 133
              :::|||:::::::|||   :::||||||:::|||   :::|||   |||      |||:::
Db          3 ATGCCCTCCGTGTACACGATTGTGTTCATTGTCAGCCTTCCTCTGAACGTCCTGGCCATC 62

Qy        134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153
                  :::::::::|||      :::   :::                |||:::      :::
Db         63 GCAGTGTTCGTCTTGAGGATGAAGGTCAAGAAG---CCGGCCGTGGTGTACATGCTGCAC 119

Qy        154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
              ||||||:::||||||   ||||||          ||||||||||||||:::|||:::
Db        120 CTGGCCATGGCCGACGTGCTCTTCGTGTCGGTGCTCCCCTTCAAGATCAGCTACTACTTC 179

Qy        174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
              :::|||   :::|||   ||||||      :::||||||      |||      ||||||
Db        180 TCCGGCACTGATTGGCAGTTCGGGTCTGGAATGTGCCGCTTCGCCACCGCAGCGTTTTAC 239

Qy        194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
              ||||||||||||   ||||||:::|||:::      |||||||||:::|||:::||||||
Db        240 GGGAACATGTACGCCTCCATCATGCTCATGACGGTCATAAGCATTGACCGGTTCCTGGCG 299

Qy        214 IleValHisPro---------PheThrTyrArgGlyLeuProLysHisThrTyrAlaLeu 230
              :::|||:::|||           ::::::|||   |||   :::      :::
Db        300 GTGGTGTATCCGATCCAGTCCCTGTCCTGGCGCACTCTGGGCCGAGCCAACTTC------ 353

Qy        231 ValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeu 250
              ||||||   ::::::|||            :::   ::::::|||      :::
Db        354 ---ACTTGCGTGGTCATTTGGGTGATGGCCATCATGGGGGTGGTGCCCCTTCTCCTCAAG 410

Qy        251 LysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspVal------His 268
              :::|||       :::      :::||||||||||||||||||||
Db        411 GAGCAGACCACCCGAGTTCCGGGACTCAACATCACCACCTGCCACGACGTCCTCAGTGAG 470

Qy        269 AsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePhe 288
              |||      :::           |||      |||||||||   :::   :::
Db        471 AACCTGATGCAAGGC---------TTTTACTCGTACTACTTCTCGGCCTTCTCCGCCATC 521

Qy        289 GlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeu 308
              ||||||:::|||   ::::::          ||||||   :::|||||||||   |||
Db        522 TTCTTTCTTGTGCCGTTGATCGTTTCCACGGTCTGCTACACGTCCATCATCCGGTGCCTG 581

Qy        309 Asn--------------------AlaTyrAspHisArgTrpLeuTrpTyrValLysAla 321
              :::                    :::      |||   |||:::
Db        582 AGCTCCTCCGCGGTTGCCAACCGGAGCAAGAAGTCGCGGGCTTTGTTC---------CTG 632

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
              |||      :::      ||||||   :::||||||   |||   |||::::::||||||
Db        633 TCTGCCGCGGTGTTCTGCATCTTCATCGTCTGCTTTGGGCCCCC-AACGTCCTCCTGATT 691
```

```
Qy        342 IleHisHisAlaAsnTyrTyrTyr-----------AsnAsnThrAspGlyLeuTyrPhe 357
              :::|||          |||   :::                  |||:::      ||||||
Db        692 GTGCAC---------TACCTTTTCCTCTCCGACAGTCCTGGTACAGAGGCAGCCTACTTT 742

Qy        358 IleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyr 377
              ||||||:::   :::|||:::   |||::::::   |||:::||||||   :::|||
Db        743 GCTTACCTCCTCTGCGTCTGTGTGAGCAGCGTGAGCTGCTGCATCGATCCGTTGATTTAC 802

Qy        378 PheLeuMetSer 381
              :::      |||
Db        803 TACTACGCCTCC 814

RESULT 11
BG924078
LOCUS       BG924078                801 bp    mRNA    linear   EST 05-JUN-2001
DEFINITION  602823635F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4952433 5',
            mRNA sequence.
ACCESSION   BG924078
VERSION     BG924078.1  GI:14304554
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10910  row: e  column: 10
            High quality sequence start: 25
            High quality sequence stop: 799.
FEATURES             Location/Qualifiers
     source          1. .801
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4952433"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      166 a    228 c    193 g    213 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              3.4e-36         Length:         801
Score:                  437.00          Matches:        107
Percent Similarity:     55.20%          Conservative:   47
Best Local Similarity:  38.35%          Mismatches:     95
Query Match:            20.26%          Indels:         32
DB:                     13              Gaps:           10

US-09-208-629F-6 (1-408) x BG924078 (1-801)

Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGlu 68
              |||      |||::::::   |||   :::   |||   :::      ||||||
Db          7 CTGTACGNCCCGAGTATTGGGATATCGTCGACCAGGCCGTCGCGCAACAACAGTAA-AGG 65

Qy         69 GluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLys 88
              :::      :::      ||||||      |||          ||||||   |||
Db         66 AAGAAGTCTTATGGCAGATTGGAA------ACCCAGCCTCCAATCACTGGGAAA---GGG 116

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
                 |||   |||           :::          :::|||::::::      ||||||
Db        117 GTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCCATC------CTCACC 170

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
                 |||:::|||      :::|||   :::|||::::::|||||||||:::|||:::
Db        171 GGGAAGCTGACCACTGTCTTTCTTCCGGTCGTCTACATTATTGTGTTTGTGATTGGTTTG 230

Qy        129 ProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThr 148
              |||:::|||   :::   ||||||:::      |||||||||:::
Db        231 CCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAGAAACACCCCGCC 290

Qy        149 ValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPhe 167
              |||   |||      |||||||||:::||||||   |||      :::      |||
Db        291 GTGATTTACATGGCCAACCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCCTG 350

Qy        168 LysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAla 187
                 ||||||||||||||:::|||||||||||||||:::||||||   ||||||:::
Db        351 GCCATTGCCTACCACCTACATGGCAACAACTGGGTCTATGGGGAGGCCCTGTGCAAGGTG 410

Qy        188 ThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSer 207
                       ||||||||||||||||||||||||||||   :::   |||:::|||
Db        411 CTCATTGGCTTTTTCTATGGCAACATGTATTGCTCCATCCTCTTCATGACCTGCCTCAGC 470

Qy        208 IleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThr 227
              :::   ||||||      ||||||:::|||          |||   |||:::
Db        471 GTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGGCACCCCAGGAAGAAG 521

Qy        228 TyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrValPheLeuTyrMet 244
                 |||   :::      |||:::          :::|||      :::||||||
Db        522 ---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATTTTTCTGGTCACC 578

Qy        245 LeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCys 264
              :::|||   :::::::::||||||      |||:::      :::||||||||||||
Db        579 ATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCCTTGAACATCACCACCTGC 638

Qy        265 HisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr------------ 280
              |||||||||                        |||   ::::::
Db        639 CACGATGTG--------------------CTGCCTGAGGAGGTATTGGTGGGGGACATG 677

Qy        281 ------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIle 298
                    ||||||:::|||||||||          ||||||   |||   :::|||
Db        678 TTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCCATCCTTAC-TGC 736

Qy        299 TyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrp 317
              ||||||      :::|||:::||||||                        ||||||
Db        737 ATCTGCTACGTGCTCATGATCAAGACGCTCCGTCTTCTG------------CTATGG 781

RESULT 12
BI394593
LOCUS       BI394593                624 bp    mRNA    linear   EST 06-AUG-2001
DEFINITION  pgp1n.pk002.l12 Normalized Chicken Pituitary/Hypothalamus/Pineal
            Library Gallus gallus cDNA clone pgp1n.pk002.l12 5' similar to
            gi|4503639 ref|NP_003941.1| coagulation factor II (thrombin)
            receptor-like 3; protease-activated receptor-4 [Homo sapiens]
            gb|AAC25699.1| (AF055917) protease-activated receptor 4 [Homo
            sapiens] gb|AAC28860.1| (AF080214) protease-activated recep, mRNA
            sequence.
ACCESSION   BI394593
VERSION     BI394593.1  GI:15087875
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Porter,T.E. and Cogburn,L.A.
  TITLE     ESTs from Normalized Chicken Pituitary/Hypothalamus/Pineal cDNA
            library USDA/IFAFS Animal Genome Project
```

```
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Gallus gallus"
                     /strain="Commercial broiler chicken"
                     /db_xref="taxon:9031"
                     /clone="pgpln.pk002.l12"
                     /clone_lib="Normalized Chicken
                     Pituitary/Hypothalamus/Pineal Library"
                     /sex="Male and Female"
                     /tissue_type="Pituitary Gland/Hypothalamus/Pineal Gland"
                     /dev_stage="Embryonic (d12,d14,d19); post-hatch (w1,w3,w5
                     ,w7,w9)"
                     /lab_host="E. Coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated from each  tissue at different
                     ages. Single pass sequencing from 5'-end"
BASE COUNT       98 a    202 c    151 g    162 t     11 others
ORIGIN

Alignment Scores:
Pred. No.:              2.7e-36        Length:        624
Score:                  436.50         Matches:       88
Percent Similarity:     58.79%         Conservative:  29
Best Local Similarity:  44.22%         Mismatches:    81
Query Match:            20.24%         Indels:        2
DB:                     13             Gaps:          1

US-09-208-629F-6 (1-408) x BI394593 (1-624)

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              |||||||||:::::|||:::|||:::   |||      ||||||:::|||      |||
Db         25 CTCGTCTTCCTGCTGGGGCTGCCCTCCACGGCCTG-GCCCTCTGGGTTCTGGCTACCAGG 83

Qy        142 ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPhe 161
              |||    :::   :::|||:::|||      |||||||||   ||||||   |||
Db         84 ACTGAGAAGCTGACCTCCACCATCTTCCTGATGAACCTGGCTGCAGCAGACCTGCTGCTC 143

Qy        162 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 181
                  ||||||   ||||||   |||:::      ||||||||||||   ||||||
Db        144 ATCTCAGTGCTGCCCCTCAAGATATTCTACTATTTCCTTGGGAACAACTGGCCCTTTGGG 203

Qy        182 GluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeu 201
              |||   |||||||||   ||||||      |||||||||||||||||||||||:::|||
Db        204 GAAGGTCTGTGCCGGGTCACCACAGCTTTCTTCTATGGGAACATGTACTGCTCAGTGCTG 263

Qy        202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
              ||||||   |||||||||:::::|||||||||||:::|||||||||||      |||
Db        264 CTGCTCACCTGCATCAGTGTGGACCGATACCTGGCGGTGGTGCATCCTTTCTTCTCGCGC 323

Qy        222 GlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
                      |||:::|||      ||||||      ||||||
Db        324 TCTTTCCGTACCCCCACCTTTGCTGCCTGGACCTGTGCTGCTGTCTGGCTCTGTGCTGCC 383

Qy        242 LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
                  ||||||      :::   :::|||   |||   |||   :::   ||||||
Db        384 GGCATCACTCTGCCCCTGACTCTGCAGCAGCAGTCGTACCCCCTGTACAGAGCAGACATC 443

Qy        262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyr 281
              |||   ||||||||||||      |||           :::      ||||||
Db        444 ACTCTGTGCCACGATGTTCTCCCTAGGCAGGAGGACGACGGG---TATTACTTCTACTAC 500

Qy        282 PheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyr 301
              |||:::   |||           ||||||   |||      ::::::::::::   |||
Db        501 TTCGTCTGCCTCATCGCCTGCGCTTTCCTTGCTCCTCTGGCAGTGATGCTGTTCAGTTAC 560

Qy        302 AlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
```

```
                  ::::::::::::                  |||:::         :::|||
Db        561 TGCTCAGTGCTACAAGCCCNNNNNNNCAGNNGGAAGCGGTACTTCNACTCCATGAAG 617

RESULT 13
AJ446632
LOCUS       AJ446632                 860 bp    mRNA    linear   EST 19-APR-2002
DEFINITION  AJ446632 riken1 Gallus gallus cDNA clone 14d16r1, mRNA sequence.
ACCESSION   AJ446632
VERSION     AJ446632.1  GI:20213853
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 860)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .860
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="14d16r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      207 a    185 c    204 g    260 t      4 others
ORIGIN

Alignment Scores:
Pred. No.:              1.18e-34       Length:        860
Score:                  423.00         Matches:       94
Percent Similarity:     54.69%         Conservative:  40
Best Local Similarity:  38.37%         Mismatches:    97
Query Match:            19.61%         Indels:        15
DB:                     9              Gaps:          6

US-09-208-629F-6 (1-408) x AJ446632 (1-860)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                        |||           :::   |||
Db         96 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 155

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                  |||||||||           |||   :::
Db        156 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 206

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                           ||||||      |||:::|||      :::|||   :::|||:::
Db        207 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 266

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::|||:::|||:::||||||:::   :::|||:::   ||||||||
Db        267 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 326

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                |||   |||      |||||||||:::||||||   |||
Db        327 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 386

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||      |||   ||||||||||||||||||||||||||||||:::|||
Db        387 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 446

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                |||||||||||||||||||||||||
Db        447 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTATGGAAATATGTACTGCTCCATT 506
```

```
Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        507 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 566

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                      |||         ||||||                ::::|||      :::
Db        567 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 623

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro-As 260
              |||       :::|||   :::::::::   |||       |||:::   :::||| ::
Db        624 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACGGTGTATATTTTCAGACCTTAA 683

Qy        260 pIleThrThrCysHisAspVal------HisAsn-----------ThrCysGluSerSe 274
              :||||||||||||||||||||         |||           ||||||   |||::
Db        684 CATCACTACCTGCCATGATGTGTTGCCTGAAAATATTTGGCTCATGACATGTTCAAGTAC 743

Qy        274 rSerProPheGlnLeuTyr------TyrPheIleSerLeuAlaPhePheGlyPheLeuIl 292
              :   |||       |||           |||:::      |||          :::|||::
Db        744 TTT-CCTCTTCACTTGCAATTNGGACTCTTTCTAATTCCAGCTGNNCATCACTGGCTGGT 802

Qy        292 eProPheValLeu 296
              :   ::::::|||
Db        803 TGCTTACATACTN 815
```

```
RESULT 14
AJ452624
LOCUS       AJ452624                800 bp    mRNA    linear   EST 22-APR-2002
DEFINITION  AJ452624 riken1 Gallus gallus cDNA clone 31k11r1, mRNA sequence.
ACCESSION   AJ452624
VERSION     AJ452624.1  GI:20262720
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 800)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .800
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="31k11r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      188 a    170 c    200 g    239 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:              1.36e-34        Length:         800
Score:                  422.00          Matches:        87
Percent Similarity:     54.42%          Conservative:   36
Best Local Similarity:  38.50%          Mismatches:     91
Query Match:            19.56%          Indels:         12
DB:                     9               Gaps:           4

US-09-208-629F-6 (1-408) x AJ452624 (1-800)

Qy         62 AlaProProAsnSerPheGluGluProPheProSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                  |||          :::   |||
Db        116 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 175

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                  ||||||||          |||   :::
Db        176 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 226

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                        ||||||      |||:::|||      :::|||   :::|||:::
Db        227 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 286

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::|||:::|||:::||||||:::   :::|||:::   |||||||||
Db        287 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 346

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                |||   |||       |||||||||:::|||||   |||
Db        347 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 406

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||       |||   ||||||||||||||||||||||||||||||:::|||
Db        407 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 466

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                ||||||||||||||||||||||||
Db        467 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 526

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        527 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 586

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                      |||         ||||||                ::::|||      :::
Db        587 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 643

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
                 :::       :::|||   :::::::::   |||       |||:::         :::
Db        644 TTGGTGGGAACCATTCCGCTGTATCTTGTCAATCAGACNGNGTATATTTCAGACCTTAAC 703

Qy        261 IleThrThrCysHisAspVal--------------------HisAsnThrCysGluSer 273
              ||||||||||||||||||||                    |||:::   :::|||
Db        704 ATCACTACCTGCCATGATGTGTTGCCTGGAAAATATTTNGCTCATGACATGTTCAAGTCT 763

Qy        274 SerSerProPheGlnLeu 279
                 |||      ||||||
Db        764 TTCTCTTACTTGCAATTG 781
```

```
RESULT 15
AJ456719
LOCUS       AJ456719                808 bp    mRNA    linear   EST 22-APR-2002
DEFINITION  AJ456719 riken1 Gallus gallus cDNA clone 9f2r1, mRNA sequence.
ACCESSION   AJ456719
VERSION     AJ456719.1  GI:20266815
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 808)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .808
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="9f2r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
```

```
BASE COUNT      189 a    169 c    204 g    244 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              1.56e-34        Length:         808
Score:                  421.50          Matches:        91
Percent Similarity:     52.30%          Conservative:   34
Best Local Similarity:  38.08%          Mismatches:     95
Query Match:            19.54%          Indels:         19
DB:                     9               Gaps:           5

US-09-208-629F-6 (1-408) x AJ456719 (1-808)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                    |||          :::   |||
Db        118 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 177

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
              :::                 |||||||||           |||   :::
Db        178 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 228

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                      ||||||      |||:::|||      :::|||   :::|||:::
Db        229 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 288

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::::|||:::|||:::|||||||:::   :::|||:::   |||||||||
Db        289 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 348

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                 |||   |||       |||||||||:::||||||   |||
Db        349 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 408

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||       |||   ||||||||||||||||||||||||||||||:::|||
Db        409 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 468

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                 ||||||||||||||||||||||||
Db        469 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 528

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        529 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 588

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                 |||          ||||||                :::|||      :::
Db        589 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 645

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              |||      :::|||   :::::::::   |||      |||:::      :::
Db        646 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACGGNGTATATTTCAGACCTTAAC 705

Qy        261 IleThrThrCysHisAspValHisAsn--------------------ThrCysGluSer 273
              ||||||||||||||||||                            ||||||   |||
Db        706 ATCACTACCTGCCATGATGGTGTTGNCTGAAAATATTTTTGGCTCATGACATGTTCAAGT 765

Qy        274 SerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              :::          |||       |||   |||       |||||||||
Db        766 ACT--------------------TTCTTTCACTTGCAATTTGGACTCTTCCTAATT 801

RESULT 16
AJ456561
LOCUS       AJ456561                890 bp    mRNA    linear   EST 22-APR-2002
DEFINITION  AJ456561 riken1 Gallus gallus cDNA clone 8o7r1, mRNA sequence.
ACCESSION   AJ456561
VERSION     AJ456561.1  GI:20266657
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 890)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .890
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="8o7r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      211 a    185 c    219 g    270 t      5 others
ORIGIN

Alignment Scores:
Pred. No.:              3.74e-34        Length:         890
Score:                  418.50          Matches:        95
Percent Similarity:     52.90%          Conservative:   42
Best Local Similarity:  36.68%          Mismatches:     107
Query Match:            19.40%          Indels:         16
DB:                     9               Gaps:           5

US-09-208-629F-6 (1-408) x AJ456561 (1-890)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                    |||          :::   |||
Db        133 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 192

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
              :::                 |||||||||           |||   :::
Db        193 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 243

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                      ||||||      |||:::|||      :::|||   :::|||:::
Db        244 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 303

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::::|||:::|||:::|||||||:::   :::|||:::   |||||||||
Db        304 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 363

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                 |||   |||       |||||||||:::||||||   |||
Db        364 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 423

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||       |||   ||||||||||||||||||||||||||||||:::|||
Db        424 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 483

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                 ||||||||||||||||||||||||
Db        484 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 543

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        544 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 603

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                 |||          ||||||                :::|||      :::
Db        604 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 660

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro-As 260
              |||      :::|||   :::::::::   |||      |||:::   :::||| ::
Db        661 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACGGTGTATATTTTCAGACCTTAA 720

Qy        260 pIleThrThrCysHisAspVal--------------------HisAsnThrCysGluSe 273
              :||||||||||||||||||||                    |||:::      :::||
```

```
Db        721 CATCACTACCTGCCATGATGTGTTGCCTGAAAATATTTTGGCTCATGACAT-GTTCAAAG 779

Qy        273 rSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIlePr 293

Db        780 TACTTTCTTCTTCACTTTGCAATTGGGACTCTTCCTTAAATTC------CAGCTTGNCAA 833

Qy        293 oPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyr 311

Db        834 TCACTGNTTGGTGCTTTACATACTTAANGGATTAAAGACTNTTGGNAATGCTTTC 888

RESULT 17
AJ453402
LOCUS       AJ453402                741 bp    mRNA    linear   EST 22-APR-2002
DEFINITION  AJ453402 riken1 Gallus gallus cDNA clone 34e17r1, mRNA sequence.
ACCESSION   AJ453402
VERSION     AJ453402.1  GI:20263498
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="34e17r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      178 a    157 c    185 g    220 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              4.72e-34        Length:         741
Score:                  416.50          Matches:        83
Percent Similarity:     56.04%          Conservative:   33
Best Local Similarity:  40.10%          Mismatches:     86
Query Match:            19.31%          Indels:         5
DB:                     9               Gaps:           3

US-09-208-629F-6 (1-408) x AJ453402 (1-741)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81

Db         96 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 155

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101

Db        156 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 206

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121

Db        207 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTCTTCCCACTGTCTATGTC 266

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141

Db        267 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 326

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160

Db        327 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 386

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180

Db        387 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 446

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200

Db        447 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATCTTTTTATGGAAATATGTACTGCTCCATT 506

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220

Db        507 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 566

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240

Db        567 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 623

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260

Db        624 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACGGTGTATATTTCANACCTTAAC 683

Qy        261 IleThrThrCysHisAspVal 267

Db        684 ATCACTACCTGCCATGATGTG 704

RESULT 18
BE146323/c
LOCUS       BE146323                265 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  MR0-HT0209-010500-110-g09 HT0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE146323
VERSION     BE146323.1  GI:8609047
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 265)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR0-HT0209-010
            500-110-g09&t3=2000-05-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 23
            High quality sequence stop: 265.
FEATURES             Location/Qualifiers
     source          1. .265
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT0209"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
```

```
                    low stringency conditions."
BASE COUNT       94 a     41 c     70 g     60 t
ORIGIN

Alignment Scores:
Pred. No.:                2.61e-34       Length:          265
Score:                    413.00         Matches:         80
Percent Similarity:       96.39%         Conservative:    0
Best Local Similarity:    96.39%         Mismatches:      2
Query Match:              19.15%         Indels:          1
DB:                       10             Gaps:            0

US-09-208-629F-6 (1-408) x BE146323 (1-265)

Qy        273 SerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 TCCTCATCTCCCTTCCAACTCTATTACTTCATCTCCTTGGCGTTCTTTGGATTCTTAATT 189

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CCATTTGTGCTTATCATCTACTGCTATGCAGCCATCATCCGGACACTTAATGCATACGAT 129

Qy        313 HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCys 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CATAGATGGTTGTGGTATGTTAAGGCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGC 69

Qy        333 PheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyr-TyrTyrAsnAsnTh 352
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         68 TGTGCTCCAAGCAATATTATTCTTATTATTCACCATGCTAACTATCTACTACAACAACAC 9

Qy        352 rAspGly 354
              |   |||
Db          8 TCATGGC 2

RESULT 19
AJ456135
LOCUS       AJ456135                845 bp    mRNA    linear   EST 22-APR-2002
DEFINITION  AJ456135 riken1 Gallus gallus cDNA clone 7k4r1, mRNA sequence.
ACCESSION   AJ456135
VERSION     AJ456135.1  GI:20266231
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 845)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .845
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="7k4r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      194 a    185 c    213 g    252 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:                1.52e-33       Length:          845
Score:                    412.50         Matches:         84
Percent Similarity:       56.25%         Conservative:    33
Best Local Similarity:    40.38%         Mismatches:      85
Query Match:              19.12%         Indels:          6
DB:                       9              Gaps:            3

US-09-208-629F-6 (1-408) x AJ456135 (1-845)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||  |||                  |||              :::    |||
Db        133 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 192

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                 |||||||||              |||   :::
Db        193 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 243

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                          ||||||      |||:::|||      :::|||   :::|||:::
Db        244 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 303

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPheArg 141
              :::||||||:::::||:::|||:::||||||:::   :::|||:::   |||||||||
Db        304 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 363

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                |||  |||      |||||||||:::||||||  |||
Db        364 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 423

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||  |||       |||  ||||||||||||||||||||||||||||:::|||
Db        424 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 483

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||  ||||||:::                 ||||||||||||||||||||||||
Db        484 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 543

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        544 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 603

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                       |||         ||||||              :::|||      :::
Db        604 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 660

Qy        241 Phe-LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAs 260
              ||| |||     :::|||   :::::::::   |||      |||:::          ::
Db        661 TTTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACNGTGTATATTTCAGACCTTAA 720

Qy        260 pIleThrThrCysHisAspVal 267
              :||||||||||||||||||||
Db        721 CATCACTACCTGCCATGATGTG 742

RESULT 20
BM439733/c
LOCUS       BM439733                602 bp    mRNA    linear   EST 01-FEB-2002
DEFINITION  pgr1n.pk001.b9 Normalized Chicken Reproductive Tract cDNA Library
            (pgr1n) Gallus gallus cDNA clone pgr1n.pk001.b9 5' similar to
            gb|AAK61908.1|AF384819_1 (AF384819) coagulation factor II
            receptor-like 3 [Homo sapiens], mRNA sequence.
ACCESSION   BM439733
VERSION     BM439733.1  GI:18470508
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   Cogburn,L.A. and Nys,Y.
  TITLE     ESTs from Normalized Chicken Reproductive Tract cDNA library-
            University of Delaware and INRA, Tours-Poultry Unit Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
```

```
          Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .602
                     /organism="Gallus gallus"
                     /strain="Commercial broiler and layer"
                     /db_xref="taxon:9031"
                     /clone="pgr1n.pk001.b9"
                     /clone_lib="Normalized Chicken Reproductive Tract cDNA
                     Library (pgr1n)"
                     /sex="Male and Female"
                     /tissue_type="Testis, ovary and oviduct"
                     /dev_stage="Various stages;embryonic, post-hatch, immature
                     and sexually-mature"
                     /lab_host="E. coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from three total
                     RNA pools from each tissue (testis 25%, ovary 25%, and
                     oviduct 50% of final RNA pool); Single pass sequencing
                     from 5'-end"
BASE COUNT      172 a    143 c    192 g     92 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:              1.97e-33       Length:         602
Score:                  409.50         Matches:        86
Percent Similarity:     49.37%         Conservative:   32
Best Local Similarity:  35.98%         Mismatches:     75
Query Match:            18.98%         Indels:         46
DB:                     13             Gaps:           3

US-09-208-629F-6 (1-408) x BM439733 (1-602)

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArg 143
              |||::::::|||:::|||:::|||   :::      |||:::|||      ||||||
Db        601 TTCCTGCTGGGGCTGCCCTCCAACGGCCTGGCNNNCTGGGTTCTGGCTACCAGGACTGAG 542

Qy        144 SerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
                :::   :::|||:::|||      |||||||||   ||||||   |||
Db        541 AAGCTGACCTCCACCATCTTCCTGATGAACCTGGCTGCAGCAGACCTGCTGCTCATCTCA 482

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
              ||||||   ||||||   |||:::      ||||||||||||   |||||||||
Db        481 GTGCTGCCCCTCAAGATATTCTACTATTTCCTTGGGAACAACTGGCCCTTTGGGGAAGGT 422

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
              |||||||||  ||||||      |||||||||||||||||||||||:::|||||||||
Db        421 CTGTGCCGGGTCACCACAGCTTTCTTCTATGGGAACATGTACTGCTCAGTGCTGCTGCTC 362

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
              |||||||||::::::||||||||||:::||||||||||       |||
Db        361 ACCTGCATCAGTGTGGACCGATACCTGGCGGTGGTGCATCCTTTCTTCTCGCGCTCTTTC 302

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
                  |||:::|||      ||||||      ||||||
Db        301 CGTACCCCCACCTTCGCTGCCTGGACCTGTGCTGCTGTCTGGCTCTGTGCTGCCGGCATC 242

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ||||||      :::   :::|||   |||   |||   :::   |||||||||
Db        241 ACTCTGCCCCTGACTCTGCAGCAGCAGTCGTACCCCCTGTACAGAGCAGACATCACTCTG 182

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||
Db        181 TGCCACGATGTT---------------------------------------------- 170

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
                        |||:::   :::|||   :::|||::::::
Db        169 ---------CTCCCTGGTTACACACTGCCAGGGCTGCTTTTGTTA-------------- 134

Qy        304 IleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeu 323
                                                     |||          |||
Db        133 ---------------------------------------GTACACCCAGGTCTG 119

Qy        324 LeuIleLeu-ValIlePheThrIleCysPheAlaProSer----------------As 337
              |||:::::: ::::::         |||:::   ||||||
```

```
Db        118 CTTTTAATGCATGCTGTGAGCTGGCTGCTATCTGCCTTCTATTCATTTTTCTGCTACTGG 59

Qy        337 nIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeu 355
                :::   ::::::   |||      |||          |||:::
Db         58 CTTCCTTGCTGTCCTTGAACACACGTTCCTTTACCTTCAGAAGGCCTCGGGGATG 4

RESULT 21
AJ446611
LOCUS       AJ446611                747 bp    mRNA    linear   EST 19-APR-2002
DEFINITION  AJ446611 riken1 Gallus gallus cDNA clone 14c16r1, mRNA sequence.
ACCESSION   AJ446611
VERSION     AJ446611.1  GI:20213832
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 747)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .747
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="14c16r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      178 a    158 c    187 g    224 t
ORIGIN

Alignment Scores:
Pred. No.:              2.67e-33       Length:         747
Score:                  409.50         Matches:        81
Percent Similarity:     55.61%         Conservative:   33
Best Local Similarity:  39.51%         Mismatches:     86
Query Match:            18.98%         Indels:         5
DB:                     9              Gaps:           3

US-09-208-629F-6 (1-408) x AJ446611 (1-747)

Qy         62 AlaProProAsnSerPheGluGluProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                          |||           :::   |||
Db         96 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 155

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                ::                    |||||||||           |||    :::
Db        156 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 206

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                          ||||||      |||:::|||      :::|||   :::|||:::
Db        207 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 266

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||::::::|||:::|||:::||||||:::   :::|||:::   |||||||||
Db        267 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 326

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                |||   |||      |||||||||:::||||||   |||
Db        327 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 386

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||      |||   |||||||||||||||||||||||||||||:::|||
Db        387 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 446
```

```
Qy      181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
Db      447 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 506

Qy      201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
Db      507 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 566

Qy      221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
Db      567 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 623

Qy      241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db      624 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACGGTGTATATTTCAGACCTTAAC 683

Qy      261 IleThrThrCysHis 265
Db      684 ATCACTACCTGCCAT 698

RESULT 22
AJ451396
LOCUS       AJ451396                773 bp    mRNA    linear   EST 22-APR-2002
DEFINITION  AJ451396 riken1 Gallus gallus cDNA clone 28j6r1, mRNA sequence.
ACCESSION   AJ451396
VERSION     AJ451396.1  GI:20261492
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .773
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="28j6r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      182 a    165 c    190 g    235 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.81e-33        Length:         773
Score:                  409.50          Matches:        84
Percent Similarity:     56.25%          Conservative:   33
Best Local Similarity:  40.38%          Mismatches:     85
Query Match:            18.98%          Indels:         6
DB:                     9               Gaps:           3

US-09-208-629F-6 (1-408) x AJ451396 (1-773)

Qy       62 AlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThrGlyAla 81
Db       97 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 156

Qy       82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db      157 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 207

Qy      102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
```

```
Db      208 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 267

Qy      122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db      268 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 327

Qy      142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
Db      328 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 387

Qy      161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
Db      388 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 447

Qy      181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
Db      448 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTATGGAAATATGTACTGCTCCATT 507

Qy      201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
Db      508 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 567

Qy      221 ArgGlyLeuProLysHisThrTyrAlaLeu-ValThrCysGlyLeuValTrpAlaThrVa 240
Db      568 ---TCAAGAAGGAAGTTTGAAATTGCCCTGGGCATCTCCCTTGCTATCTTGGATACTGAT 624

Qy      240 lPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAs 260
Db      625 TTTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACNGTGTATATTTCAGACCTTAA 684

Qy      260 pIleThrThrCysHisAspVal 267
Db      685 CATCACTACCTGCCATGATGTG 706

RESULT 23
AJ447002
LOCUS       AJ447002                749 bp    mRNA    linear   EST 19-APR-2002
DEFINITION  AJ447002 riken1 Gallus gallus cDNA clone 15f3r1, mRNA sequence.
ACCESSION   AJ447002
VERSION     AJ447002.1  GI:20214223
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="15f3r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      176 a    156 c    196 g    220 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              3.43e-33        Length:         749
Score:                  408.50          Matches:        82
Percent Similarity:     56.31%          Conservative:   34
Best Local Similarity:  39.81%          Mismatches:     84
Query Match:            18.94%          Indels:         6
```

```
DB:                 9             Gaps:         3

US-09-208-629F-6 (1-408) x AJ447002 (1-749)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                          |||           :::   |||
Db        118 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 177

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                  |||||||||                  |||   :::
Db        178 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 228

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                       ||||||      |||:::|||      :::|||   :::|||:::
Db        229 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 288

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::|||:::|||:::||||||:::   :::|||:::   |||||||||
Db        289 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 348

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                  |||   |||       |||||||||:::||||||   |||
Db        349 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 408

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||      |||   |||||||||||||||||||||||||||||||||:::|||
Db        409 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 468

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                  |||||||||||||||||||||||
Db        469 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 528

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        529 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 588

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                       |||          ||||||                  :::|||      :::
Db        589 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 645

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro-As 260
                 |||      :::|||   :::::::::   |||      |||:::   :::||| ::
Db        646 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACGGNGTATATTTTCAGACCTTAA 705

Qy        260 pIleThrThrCysHis 265
              :|||||||||||||||
Db        706 CATCACTACCTGCCAT 721

RESULT 24
AJ450835
LOCUS       AJ450835               766 bp    mRNA    linear   EST 19-APR-2002
DEFINITION  AJ450835 riken1 Gallus gallus cDNA clone 26115r1, mRNA sequence.
ACCESSION   AJ450835
VERSION     AJ450835.1  GI:20218056
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 766)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .766
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="26115r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      181 a    160 c    197 g    227 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              7.41e-33        Length:         766
Score:                  405.50          Matches:        84
Percent Similarity:     56.52%          Conservative:   33
Best Local Similarity:  40.58%          Mismatches:     85
Query Match:            18.80%          Indels:         6
DB:                     9               Gaps:           3

US-09-208-629F-6 (1-408) x AJ450835 (1-766)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                          |||           :::   |||
Db        118 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 177

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                  |||||||||                  |||   :::
Db        178 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 228

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                       ||||||      |||:::|||      :::|||   :::|||:::
Db        229 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 288

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::|||:::|||:::||||||:::   :::|||:::   |||||||||
Db        289 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 348

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                  |||   |||       |||||||||:::||||||   |||
Db        349 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 408

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||      |||   |||||||||||||||||||||||||||||||||:::|||
Db        409 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 468

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                  |||||||||||||||||||||||
Db        469 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 528

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        529 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 588

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                       |||          ||||||                  :::|||      :::
Db        589 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 645

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
                 |||      :::|||   :::::::::   |||      |||:::   :::|||
Db        646 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACNGTGTATATTTC-AGACCTTAC 704

Qy        261 IleThrThrCysHisAspVal 267
              ||||||||||||||||||||
Db        705 ATCACTACCTGCCATGATGTG 725

RESULT 25
AJ450921
LOCUS       AJ450921               740 bp    mRNA    linear   EST 19-APR-2002
DEFINITION  AJ450921 riken1 Gallus gallus cDNA clone 27a5r1, mRNA sequence.
ACCESSION   AJ450921
VERSION     AJ450921.1  GI:20218142
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
```

```
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 740)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .740
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="27a5r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      176 a    158 c    184 g    219 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:              1.89e-32        Length:         740
Score:                  401.50          Matches:        82
Percent Similarity:     55.29%          Conservative:   33
Best Local Similarity:  39.42%          Mismatches:     87
Query Match:            18.61%          Indels:         6
DB:                     9               Gaps:           3

US-09-208-629F-6 (1-408) x AJ450921 (1-740)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                         |||             :::   |||
Db         97 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 156

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                 |||||||||                 |||   :::
Db        157 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 207

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                      ||||||      |||:::|||      :::|||   :::|||:::
Db        208 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 267

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::|||:::|||:::|||||||:::   :::|||:::   |||||||||
Db        268 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 327

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                 |||   |||      |||||||||:::||||||   |||
Db        328 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 387

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||      |||   |||||||||||||||||||||||||||||||:::|||
Db        388 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 447

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                 |||||||||||||||||||||||
Db        448 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 507

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        508 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 567

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                    |||      ||||||                 :::|||   :::
Db        568 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 624

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
                        :::|||   :::::::::   |||      |||:::      :::
Db        625 TTGNTGGGAACCATTCCGCTGTATCTTGTCAATCAGACNGNGTATATTTCAGACCTTAAC 684

Qy        261 IleThrThr-CysHisAspVal 267
```

```
                 |||||||| ||||||||||||
Db        685 ATCACTACCCTGCCATGATGTG 706

RESULT 26
AJ456491
LOCUS       AJ456491                760 bp    mRNA    linear   EST 22-APR-2002
DEFINITION  AJ456491 riken1 Gallus gallus cDNA clone 816r1, mRNA sequence.
ACCESSION   AJ456491
VERSION     AJ456491.1  GI:20266587
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .760
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="816r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      179 a    165 c    197 g    217 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              1.96e-32        Length:         760
Score:                  401.50          Matches:        80
Percent Similarity:     55.56%          Conservative:   35
Best Local Similarity:  38.65%          Mismatches:     87
Query Match:            18.61%          Indels:         5
DB:                     9               Gaps:           3

US-09-208-629F-6 (1-408) x AJ456491 (1-760)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                         |||             :::   |||
Db        141 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 200

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                 |||||||||                 |||   :::
Db        201 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 251

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                      ||||||      |||:::|||      :::|||   :::|||:::
Db        252 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 311

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::|||:::|||:::|||||||:::   :::|||:::   |||||||||
Db        312 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 371

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                 |||   |||      |||||||||:::||||||   |||
Db        372 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 431

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||      |||   |||||||||||||||||||||||||||||||:::|||
Db        432 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 491

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                 |||||||||||||||||||||||
Db        492 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGAAAATATGTACTGCTCCATT 551
```

```
Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::::::::   ||||||      :::|||:::|||      :::
Db        552 CTTTTTATGACATGTCTCAATGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 611

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                      |||         ||||||               :::|||      :::
Db        612 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 668

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              |||       :::|||   :::::::::   |||       ::::::      :::
Db        669 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACNGGGNGTATTTTAGACCTTAAC 728

Qy        261 IleThrThrCysHisAspVal 267
              ||||||||||||||||||||
Db        729 ATCACTACCTGCCATGATGTG 749

RESULT 27
AJ455912
LOCUS       AJ455912                730 bp    mRNA    linear   EST 22-APR-2002
DEFINITION  AJ455912 riken1 Gallus gallus cDNA clone 6p18r1, mRNA sequence.
ACCESSION   AJ455912
VERSION     AJ455912.1  GI:20266008
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="6p18r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      177 a    156 c    179 g    217 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.37e-32        Length:         730
Score:                  400.50          Matches:        84
Percent Similarity:     56.52%          Conservative:   33
Best Local Similarity:  40.58%          Mismatches:     85
Query Match:            18.57%          Indels:         6
DB:                     9               Gaps:           3

US-09-208-629F-6 (1-408) x AJ455912 (1-730)

Qy         62 AlaProProAsnSerPheGluGluProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||  |||                |||           :::  |||
Db         87 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 146

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
              :::                 |||||||||             |||  :::
Db        147 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 197

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                      ||||||      |||:::|||      :::|||   :::|||:::
Db        198 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 257

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
                 :::||||||:::::|||:::|||:::||||||:::   :::|||:::   |||||||||
Db        258 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 317

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                  |||   |||      ||||||||:::||||||   |||
Db        318 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 377

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||       |||   |||||||||||||||||||||||||||||:::|||
Db        378 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 437

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::           ||||||||||||||||||||||||||||||||
Db        438 GGGGAAGGTCTCTGCAAAGTGCTTATTGG-ATTTTTTATGGAAATATGTACTGCTCCATT 496

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        497 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 556

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                      |||         ||||||               :::|||      :::
Db        557 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 613

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              |||       :::|||   :::::::::   |||       |||:::       :::
Db        614 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACNGGGTATATTTCAAACCTTAAC 673

Qy        261 IleThrThrCysHisAspVal 267
              ||||||||||||||||||||
Db        674 ATCACTACCTGCCATGATGTG 694

RESULT 28
CNS03DLV/c
LOCUS       CNS03DLV                817 bp    DNA     linear   GSS 15-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            017G16 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL239260
VERSION     AL239260.1  GI:7898395
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 817)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 817)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 817)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .817
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="017G16"
```

```
                    /clone_lib="G"
                    /note="Genoscope sequence ID : C0BG017BD08LP1~end : T7"
BASE COUNT      218 a    183 c    232 g    167 t     17 others
ORIGIN

Alignment Scores:
Pred. No.:                2.54e-31       Length:          817
Score:                    391.50         Matches:         76
Percent Similarity:       60.67%         Conservative:    32
Best Local Similarity:    42.70%         Mismatches:      66
Query Match:              18.15%         Indels:          4
DB:                       17             Gaps:            2

US-09-208-629F-6 (1-408) x CNS03DLV (1-817)

Qy          92 GluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeu 111
               ||||||          :::   |||:::::          |||   |||         |||
Db         560 GAATCCAGGGATGGCGTTAGCGTGGAATCGTACGGTAGAGGAACTCTCTGCGGCCGCCTC 501

Qy         112 SerThrLysLeuIleProAlaIle-TyrLeuLeuValPheValValGlyValProAlaAs 131
               :::|||         |||   ::: |||:::   ||||||||||||||:::|||   ||
Db         500 ACCACCGTCTTCTTCCCCGTCGTCATACATCACCGTCTTCGTGGTGGGCCTCCCCACCAA 441

Qy         131 nAlaValThrLeuTrpMetLeuPhePheArgThr---ArgSerIleCysThrThrValPh 150
               ||||:::   :::|||:::      |||||||||   :::          :::::::::||
Db         440 CGCCCTGGCCATCTGGGTCTTCCTCTTCCGCACCAAGAAGAAGCATCCATCCTCCATCTT 381

Qy         150 eTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAl 170
               |      |||||||||:::||||||    ||||||   :::   :::|||    ||||||||
Db         380 CATGGCCAACCTGGCCGTGGCCGACCTGCTCTTCGTCATCTGGGTCCCGCTGAAAATCGC 321

Qy         170 aTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrVa 190
               ||||||||||||||||||:::||||||:::||||||    ||||||:::
Db         320 CTACCACCTCAACGGGAACGACTGGGTCTACGGCGAGTCCCTGTGCAAGGTCCTGGTGGC 261

Qy         190 lIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnAr 210
               : ||||||||||||||||||||||||||      :::|||||||||||||:::    ||
Db         260 CTTCTTCTACGGCAACATGTACTGCTCCATCGCCTTCATCGCGTGCATCAGCGTCCAGCG 201

Qy         210 gTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLe 230
               |      |||:::::||||||          ||||||                  |||::
Db         200 CTGCAAGGCCGTGCTCCACCCGCTCGACAAGAGGGGG------GGCAGCGGTGCGGCCGT 147

Qy         230 uValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLe 250
               :          :::|||      |||:::|||      :::|||   ::::::
Db         146 GGGCGTGTCCGCGGCCATCTGGCTGCTGGTGTGGCTCATCACCGTTCCTCTGTACCTGTA 87

Qy         250 uLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspVal 267
                  |||:::      :::            |||   ||||||||||||||
Db          86 CGATCAGCAGGTCCTGGTGACAAACCTGGGCATCCGCACCTGCCACGACGTC 35

RESULT 29
BC013202
LOCUS       BC013202               2020 bp    mRNA    linear   HTC 29-AUG-2001
DEFINITION  Homo sapiens, clone IMAGE:4214482, mRNA.
ACCESSION   BC013202
VERSION     BC013202.1  GI:15342000
KEYWORDS    HTC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2020)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalon@bcm.tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,
            Muzny,D.M., Gibbs,R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 26 Row: h Column: 23
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 2695873
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1. .2020
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4214482"
                     /tissue_type="Brain, anaplastic oligodendroglioma with
                     1p/19q loss"
                     /clone_lib="NCI_CGAP_Brn67"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
BASE COUNT      410 a    689 c    508 g    413 t
ORIGIN

Alignment Scores:
Pred. No.:                1.49e-30       Length:          2020
Score:                    389.50         Matches:         107
Percent Similarity:       44.54%         Conservative:    48
Best Local Similarity:    30.75%         Mismatches:      132
Query Match:              18.06%         Indels:          62
DB:                       11             Gaps:            12

US-09-208-629F-6 (1-408) x BC013202 (1-2020)

Qy          62 AlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThrGlyAla 81
               |||||||||              |||      :::      |||
Db         122 GCTCCCCCAGGTCTGATCACCAACTTCTCCCTGGCCACGGCAGAG--------------- 166

Qy          82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                                  :::|||   :::|||
Db         167 ------------------CAATGTGGCCAGGAG-------------------------- 181

Qy         102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                                        :::   |||          |||      :::   ||||||
Db         182 ---------------------ACGCCACTGGAGAACATGCTGTTCGCCTCCTTCTACCTT 220

Qy         122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
               |||   |||::::::  :::        |||   :::   ||||||   ||||||   |||
Db         221 CTGGATTTTATCCTGGCTTTAGTTGGCAATACCCTGGCTCTGTGG---CTTTTCATCCGA 277

Qy         142 ThrArgSerIleCysThr------ThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
                           |||         ||||||      :::||||||:::||||||
Db         278 GACCACAAGTCCGGGACCCCGGCCAACGTGTTCCTGATGCATCTGGCCGTGGCCGACTTG 337

Qy         160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
                     :::   ||||||   ::::::   ||||||   :::|||||:::|||
Db         338 TCGTGCGTGCTGGTCCTGCCCACCCGCCTGGTCTACCACTTCTCTGGGAACCACTGGCCA 397

Qy         180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
               |||||||||:::   ||||||  |||      :::||||||   |||||||||   |||
Db         398 TTTGGGGAAATCGCATGCCGTCTCACCGGCTTCCTCTTCTACCTCAACATGTACGCCAGC 457

Qy         200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
               |||      |||  |||||||||   :::|||:::||||||||||||||||||||
Db         458 ATCTACTTCCTCACCTGCATCAGCGCCGACCGTTTCCTGGCCATTGTGCACCCGGTCAAG 517

Qy         220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
```

```
Db        518 TCCCTCAAGCTCCGCAGGCCCCTCTACGCACACCTGGCCTGTGCCTTCCTGTGGGTGGTG 577

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259

Db        578 GTGGCTGTGGC-ATGGCCCCGCTGCTGGTGAGCCCACAG---------------------- 615

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279

Db        616 ACCGTGCAGACCAACCACACGGTG------GTCTGC--------------CTGCAGCTG 654

Qy        280 Tyr----------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292

Db        655 TACCGGGAGAAGGCCTCCCACCATGCCCTGGTGTCCCTGGCA---GTGGCCTTCACCTTC 711

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr------------Leu 308

Db        712 CCGTTCATCACCACGGTCACCTGCTACCTGCTGATCATCCGCAGCCTGCGGCAGGGCCTG 771

Qy        309 AsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328

Db        772 CGTGTGGAGAAGCGCCTCAAGACCAAGGCAGTGCGCATGATCGCCATAGTGCTGGCCATC 831

Qy        329 PheThrIleCysPheAlaProSerAsnIle---------IleLeuIleIleHisHisAla 345

Db        832 TTCCTGGTCTGCTTCGTGCCCTACCACGTCAACCGCTCCGTCTACGTGCTGCACTACCGC 891

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGly------LeuTyrPheIleTyrLeuIleAlaLeu 363

Db        892 AGCCATGGGGCCTCCTGCGCCACCCAGCGCATCCTGGCCCTGGCAAACCGCATCACCTCC 951

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383

Db        952 TGCCTCACCAGCCTCAACGGGGCACTCGACCCCATCATGTATTTCTTCGTGGCTGAGAAG 1011

Qy        384 ArgAsnHisSerThrAlaTyrLeu 391

Db       1012 TTCCGCCACGCCCTGTGCAACTTG 1035
```

RESULT 30
AJ450259
LOCUS AJ450259 746 bp mRNA linear EST 19-APR-2002
DEFINITION AJ450259 riken1 Gallus gallus cDNA clone 24k5r1, mRNA sequence.
ACCESSION AJ450259
VERSION AJ450259.1 GI:20217480
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 746)
AUTHORS Buerstedde,J.M.
TITLE Gallus gallus bursal lymphocyte EST
JOURNAL Unpublished (2002)
COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES Location/Qualifiers
source 1. .746
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="24k5r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/note="CB inbred strain"
BASE COUNT 176 a 155 c 191 g 220 t 4 others
ORIGIN

Alignment Scores:

```
Pred. No.:              1.6e-30        Length:          746
Score:                  383.50         Matches:         79
Percent Similarity:     55.12%         Conservative:    34
Best Local Similarity:  38.54%         Mismatches:      86
Query Match:            17.78%         Indels:          6
DB:                     9              Gaps:            3

US-09-208-629F-6 (1-408) x AJ450259 (1-746)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81

Db        118 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 177

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101

Db        178 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 228

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121

Db        229 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 288

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141

Db        289 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 348

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160

Db        349 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 408

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180

Db        409 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 468

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200

Db        469 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 528

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220

Db        529 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 588

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240

Db        589 ---TCAAGAANGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 645

Qy        241 PheLeuTyrMetLeu-ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAs 260

Db        646 TTGTTGGGAACCATTCCCGCTGTATCTTGTCAATCAAACNGNGTNTATTTTAAACCTTAA 705

Qy        260 pIleThrThrCys 264

Db        706 CATCACTACCTGC 718
```

RESULT 31
AJ451155
LOCUS AJ451155 798 bp mRNA linear EST 19-APR-2002
DEFINITION AJ451155 riken1 Gallus gallus cDNA clone 27m8r1, mRNA sequence.
ACCESSION AJ451155
VERSION AJ451155.1 GI:20218376
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 798)
AUTHORS Buerstedde,J.M.
TITLE Gallus gallus bursal lymphocyte EST
JOURNAL Unpublished (2002)
COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany

```
          Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .798
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="27m8r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      187 a    170 c    199 g    236 t      6 others
ORIGIN

Alignment Scores:
Pred. No.:              1.76e-30        Length:         798
Score:                  383.50          Matches:        81
Percent Similarity:     55.02%          Conservative:   34
Best Local Similarity:  38.76%          Mismatches:     87
Query Match:            17.78%          Indels:         7
DB:                     9               Gaps:           3

US-09-208-629F-6 (1-408) x AJ451155 (1-798)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                          |||             :::   |||
Db        115 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 174

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                       |||||||||              |||   :::
Db        175 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 225

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                       ||||||    . |||:::|||      :::|||   :::|||:::
Db        226 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 285

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::::|||:::|||:::|||||||:::   :::|||:::   |||||||||
Db        286 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 345

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                 |||   |||       |||||||||:::||||||   |||
Db        346 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 405

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||       |||   |||||||||||||||||||||||||||||:::|||
Db        406 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 465

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                 |||||||||||||||||||||||||
Db        466 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCATT 525

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::|||:::   ||||||      :::|||:::|||      :::
Db        526 CTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGNTGTAGTGAACCCCATAGTGCAC 585

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                       |||       ||||||              :::|||   :::
Db        586 ---TCAAGAANGAAGTCTGAAATTGCCCTGGGCATCTNCCTTGCTATCTGGATACTGATT 642

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro-As 260
                 :::      :::|||   :::::::::            |||:::          ::
Db        643 TTGGTGGGAACCATTCCGCTGTATCTTGTCAATCANACNGNGTATATTTCAGACCTTTAA 702

Qy        260 pIleThrThr-CysHisAspVal 267
              :|||||||| ||||||||||
Db        703 CATCACTACCCTGCCATGATGTG 725

RESULT 32
AJ447835
LOCUS       AJ447835               724 bp    mRNA    linear   EST 19-APR-2002
DEFINITION  AJ447835 riken1 Gallus gallus cDNA clone 17n11r1, mRNA sequence.
ACCESSION   AJ447835
VERSION     AJ447835.1  GI:20215056
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="17n11r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      168 a    155 c    190 g    210 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              6.72e-30        Length:         724
Score:                  377.50          Matches:        79
Percent Similarity:     54.90%          Conservative:   33
Best Local Similarity:  38.73%          Mismatches:     86
Query Match:            17.50%          Indels:         6
DB:                     9               Gaps:           3

US-09-208-629F-6 (1-408) x AJ447835 (1-724)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||   |||                          |||             :::   |||
Db        115 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 174

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                       |||||||||              |||   :::
Db        175 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 225

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                       ||||||      |||:::|||      :::|||   :::|||:::
Db        226 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 285

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet-LeuPhePheAr 141
              :::||||||:::::::|||:::|||:::||||||:::   :::|||:::    ||||||||
Db        286 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTTCAG 345

Qy        141 gThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLe 160
              ||||:::                 |||   |||       |||||||||:::||||||   ||
Db        346 AACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCT 405

Qy        160 uPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPh 180
              ||||   |||       |||   ||||||||||||||||||||||||||||||:::||
Db        406 CTTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTT 465

Qy        180 eGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIl 200
              |||||||   ||||||:::                 |||||||||||||||||||||||
Db        466 TGGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTTATGGAAATATGTACTGCTCCAT 525

Qy        200 eLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTy 220
              ||||   :::   |||:::|||:::   ||||||      :::|||:::|||      ::
Db        526 TCTTTTTATGACATGTCTCAGTGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCA 585

Qy        220 rArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVa 240
              :                 |||       ||||||              :::|||      ::
Db        586 C---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGAT 642
```

```
Qy        240 lPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAs 260
              :   |||     :::|||   :::::::::   |||     |||:::         ::
Db        643 TTTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAGACNGGGTATATTTCAGACCTTAA 702

Qy        260 pIleThrThr 263
              :|||||||||
Db        703 CATCACTACC 712

RESULT 33
BJ068617
LOCUS       BJ068617                633 bp    mRNA    linear   EST 11-DEC-2001
DEFINITION  BJ068617 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL107d22 5', mRNA sequence.
ACCESSION   BJ068617
VERSION     BJ068617.1  GI:17494676
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 633)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .633
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL107d22"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT      154 a    154 c    143 g    182 t
ORIGIN

Alignment Scores:
Pred. No.:              8.04e-30        Length:         633
Score:                  376.00          Matches:        78
Percent Similarity:     55.22%          Conservative:   33
Best Local Similarity:  38.81%          Mismatches:     65
Query Match:            17.43%          Indels:         25
DB:                     13              Gaps:           5

US-09-208-629F-6 (1-408) x BJ068617 (1-633)

Qy         33 LeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysPro 52
              |||   ||||||                  |||                 ||||||
Db         69 TTATTCCCGACTACAAGCCCGGGCTCCAGAGAGAGAGGACATGGACGGCTGGCGGTGGAG 128

Qy         53 ThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPhe 72
              :::         :::   |||             |||||||||             ||||||
Db        129 TCTGCTGATACTCGGCGCTTT---GCTGTGCTTCCCAATAGTAGCAGGACAGAACCTTTC 185

Qy         73 SerAlaLeuGlu--------------------GlyTrpThrGlyAlaThrIleThrVal 85
                 ::::::|||                  |||:::|||
Db        186 TACTCCATTGAAAAAAGGAAGAGTTTTCATAGCCCATGGAGTGGA---------------- 230

Qy         86 LysIleLysCysProGluGluSerAlaSerHis-----------LeuHis-ValLysAs 101
                                  :::         |||             ||||||  |||
Db        231 -----------------GACGGGCAATGGCACTGGGAAATACTCCTACACTGTTACTCC 272

Qy        101 nAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLe 121
                          |||:::      |||:::|||      :::||||||:::|||::
Db        273 CTTTGTTTCAAAAGTGCTCTCCGATCCGCTGACCACCGCATTCTTACCTGCAGTCTACAT 332

Qy        121 uLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheAr 141
              ::::||||||::::::|||:::|||:::||||||:::   ||||||:::   ||||||||
Db        333 TATAGTGTTCATCATTGGCTTACCAAGCAACGCCATTGCACTGTGGGTCTTTTTTTTCCG 392

Qy        141 gThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLe 160
              ||||:::             ::::::   |||       |||||||||:::||||||   ::
Db        393 GACAAAGAAGAAGCACCCTTCCATGATTTATATGGCCAACCTGGCGCTGGCAGATCTGAT 452

Qy        160 uPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPh 180
              :|||   :::   ||||||   ||||||||||||||||||||||||||||||||:::::
Db        453 GTTTGTCATATGGCTCCCATTAAAGATTGCATATCATCTGAATGGAAACAATTGGATTTA 512

Qy        180 eGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIl 200
              :||||||   ||||||:::                 ||||||||||||||||||||||
Db        513 TGGAGAAGCTTTATGCAAAGTTTTGATTGGGTTCTTTTATGGGAACATGTATTGCTCCAT 572

Qy        200 eLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTy 220
              ||||   :::   |||:::|||:::   ||||||      ||||||:::|||   :::::
Db        573 TCTTTTCATGACATGCCTTAGTGTTCAGAGATACTGGGTCATCGTGAACCCCATATCCCA 632

Qy        220 r 220
              :
Db        633 C 633

RESULT 34
BI984999
LOCUS       BI984999                591 bp    mRNA    linear   EST 24-OCT-2001
DEFINITION  fu13d01.y3 Campbell zebrafish ovary Danio rerio cDNA clone 5305009
            5' similar to SW:PAR2_MOUSE P55086 PROTEINASE ACTIVATED RECEPTOR 2
            PRECURSOR ;, mRNA sequence.
ACCESSION   BI984999
VERSION     BI984999.1  GI:16376975
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library constructed by Library constructed by Invitrogen and
            donated by R. Campbell (Marine Biology Laboratory, Woods Hole, MA).
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            High quality sequence stop: 441.
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="5305009"
                     /clone_lib="Campbell zebrafish ovary"
                     /sex="female"
```

```
                     /dev_stage="4-5 months, 1 year and 2 years"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary(pooled); Vector: pCMV-SPORT6; Site_1:
                     NotI; Site_2: SalI; Whole ovaries collected from zebrafish
                     aged 4-5 months, 1 year and 2 years.  Oligo-dT primed,
                     directionally cloned.  Average insert size 2 kb.  Library
                     constructed by Invitrogen and donated by R. Campbell
                     (Marine Biology Laboratory, Woods Hole, MA)."
BASE COUNT      129 a    142 c    142 g    178 t
ORIGIN

Alignment Scores:
Pred. No.:              1.35e-29       Length:        591
Score:                  373.50         Matches:       75
Percent Similarity:     56.38%         Conservative:  31
Best Local Similarity:  39.89%         Mismatches:    59
Query Match:            17.32%         Indels:        23
DB:                     13             Gaps:          5

US-09-208-629F-6 (1-408) x BI984999 (1-591)

Qy         77 GlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHis 96
              |||:::||||||                                  ||||||:::
Db         88 GGCTTTACTGGT----------------------------GAAGAAACTAAGGAAGGT 117

Qy         97 LeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIle 116
              :::   |||   :::      :::|||  |||   |||   |||:::
Db        118 GTATCGGTGACCTCCACTGCTCTTGGTGTATTGAACAGCAAGCTCACGCAGATCTTCTTC 177

Qy        117 ProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrp 136
              |||   :::|||:::::::||||||   ||||||:::|||   ||||||:::   :::|||
Db        178 CCCGTGCTCTACATCATCGTCTTCAGTGTGGGCTTGCCCACCAATGCCATGGCCATCTGG 237

Qy        137 MetLeuPhePheArgThr---ArgSerIleCysThrThrValPheTyrThrAsnLeuAla 155
              :::      |||||||||   :::         :::::::::|||      |||||||||
Db        238 GTGTTTCTCTTTAGAACCAAGAAGAAGCACCCATCATCCATTTTTATGGCCAACCTAGCA 297

Qy        156 IleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGly 175
              :::||||||   ||||||   :::   :::|||   ||||||||||||||   ||||||
Db        298 CTGGCCGACCTTCTCTTTGTAATCTGGATCCCTTTGAAAATTGCTTATCATTTTAATGGG 357

Qy        176 AsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsn 195
              |||:::|||:::|||||||||   ||||||:::                |||||||||||
Db        358 AATCATTGGATCTTCGGAGAAGCCTTGTGTAAAGTGCTAGTAGGATTTTTCTACGGCAAT 417

Qy        196 MetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleVal 215
              ||||||||||||         :::||||||||||:::   ||||||   |||||||||
Db        418 ATGTACTGTTCCACAGCTTTCATCGCATGCATTAGCGTCCAGAGGTACTGGGCCATTGTG 477

Qy        216 HisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGly--- 234
              ||||||      :::      |||          |||      |||
Db        478 CATCCACTTTTCCAGCAG---------AAGAGGAACAACAGGTTGGCGGTGTGTGTGTCA 528

Qy        235 ---LeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
                 ||||||     |||                    :::   |||   :::
Db        529 TTGTGCGTGTGGCTGGTCGTG--------------------TGGGCTTTAACCCGGGCT 567

Qy        254 TyrTyrLeuValGlnProAspIle 261
                 |||   ::::::||||||:::
Db        568 CTGTATCCTTTACGACCAGACGTG 591

RESULT 35
AJ453748
LOCUS       AJ453748                682 bp    mRNA    linear   EST 22-APR-2002
DEFINITION  AJ453748 riken1 Gallus gallus cDNA clone 35g7r1, mRNA sequence.
ACCESSION   AJ453748
VERSION     AJ453748.1  GI:20263844
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   Buerstedde,J.M.
  TITLE     Gallus gallus bursal lymphocyte EST
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="35g7r1"
                     /clone_lib="riken1"
                     /cell_type="bursal lymphocyte"
                     /dev_stage="2-3 weeks old"
                     /note="CB inbred strain"
BASE COUNT      160 a    144 c    181 g    197 t
ORIGIN

Alignment Scores:
Pred. No.:              2.11e-29       Length:        682
Score:                  372.50         Matches:       74
Percent Similarity:     55.21%         Conservative:  32
Best Local Similarity:  38.54%         Mismatches:    81
Query Match:            17.27%         Indels:        5
DB:                     9              Gaps:          3

US-09-208-629F-6 (1-408) x AJ453748 (1-682)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||  |||                              |||          :::   |||
Db        117 GCCGCTCCCACAGAGACAAGTAGAACCAGCAGTTCAAAAGGAAGAAGTTTTGCTGGCCAG 176

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                 :::                 |||||||||              |||   :::
Db        177 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA---------TACAAAGTGGATGAC 227

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                          ||||||      |||:::|||       :::|||   :::|||:::
Db        228 TTTGCAGCAAAAGTCCTTACAGGAAAACTAACTACAGTTTTTCTTCCCACTGTCTATGTC 287

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              :::||||||:::::|||:::|||:::||||||:::   :::|||:::   ||||||||
Db        288 ATTGTCTTTATCATTGGTTTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGA 347

Qy        142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
              |||:::                |||   |||      |||||||||:::||||||   |||
Db        348 ACAAAGAAGAAACATCCTGCTGTGATTTATATGGTTAACTTGGCATTGGCAGACCTTCTC 407

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              |||   |||      |||   ||||||||||||||||||||||||||||||:::|||
Db        408 TTCGTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT 467

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||   ||||||:::                ||||||||||||||||||||||||
Db        468 GGGGAAGGTCTCTGCAAAGTGCTTGTTGGATTTTTTATGGAAATATGTACTGCTCCATT 527

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              |||   :::   |||:::::::::   ||||||      :::|||:::|||      :::
Db        528 CTTTTTATGACATGTCTCAATGTGCAAAGGTATTGGGTTGTAGTGAACCCCATAGTGCAC 587

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
                       |||        ||||||                :::|||     :::
Db        588 ---TCAAGAAGGAAGTCTGAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACTGATT 644

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGln 252
              |||      :::|||   :::::::::   |||
Db        645 TTGTTGGGAACCATTCCGCTGTATCTTGTCAATCAG 680

RESULT 36
BE005953
```

```
LOCUS       BE005953                668 bp    mRNA    linear   EST 05-JUN-2000
DEFINITION  RC0-BN0121-210300-031-d04 BN0121 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE005953
VERSION     BE005953.1  GI:8266186
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   Dias Neto,E., Gárcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC0-BN0121-210
            300-031-d04&t3=2000-03-21&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 35
            High quality sequence stop: 667.
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BN0121"
                     /dev_stage="Adult"
                     /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      137 a    172 c    157 g    202 t
ORIGIN

Alignment Scores:
Pred. No.:              3.36e-29        Length:         668
Score:                  370.50          Matches:        82
Percent Similarity:     53.21%          Conservative:   34
Best Local Similarity:  37.61%          Mismatches:     80
Query Match:            17.18%          Indels:         22
DB:                     10              Gaps:           4

US-09-208-629F-6 (1-408) x BE005953 (1-668)

Qy          61 GlyAlaProProAsnSerPheGluGluProPheSerAlaLeuGluGly--------- 77
               ||||||   |||   |||                              ::::::|||
Db           6 GGGGCGCATCCTGCTAGCAGCCTCTCTCTCTGCAGTGGCACCATCCAAGGTTGATGGCAC 65

Qy          78 ------------Trp-ThrGlyAlaThrIleThrValLysIleLysCysProGluGluSe 93
                         |||    |||   |||:::                        |||   ||
Db          66 ATCCCACGTCACTGGAAAAGGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTC 125

Qy          93 rAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerTh 113
               |||||||                                |||||      |||:::||
Db         126 TGCATCT----------------------------GTCCTCACTGGAAAACTGACCAC 155

Qy         113 rLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaVa 133
               |      :::|||   :::|||   :::||||||||||||||:::|||:::|||   ::
Db         156 TGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCAT 215

Qy         133 lThr-LeuTrpMet-LeuPhePheArgThrArgSerIleCysThrThrValPheTyr--- 151
               :   :::|||:::   |||   ||||||:::               |||   |||
Db         216 GGCCCGTGTGGGTCTTTCTTTTGCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATG 275

Qy         152 ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyr 171
               ||||||||:::|||||   |||   :::       |||   ||||||||||
Db         276 GCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTAT 335

Qy         172 HisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIle 191
               |||::::::|||||||||||:::::|||||   ||||||
Db         336 CACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTT 395

Qy         192 PheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyr 211
               |||||||||||||||||||||||||||   :::   |||:::|||:::   ||||||
Db         396 TTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTAT 455

Qy         212 LeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuVal 231
               |||||:::|||      :::          |||          |||:::
Db         456 TGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGC 512

Qy         232 ThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLys 251
                     :::|||      :::   |||      :::|||   :::::::::|||
Db         513 ATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAG 572

Qy         252 GlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspVal 267
               |||      ::::::         :::||||||||||||||||||||
Db         573 CAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGATGTT 620

RESULT 37
BF426487
LOCUS       BF426487                531 bp    mRNA    linear   EST 31-MAY-2002
DEFINITION  df70d03.y1 Xenopus laevis unfertilized egg cDNA library Xenopus
            laevis cDNA clone XENOPUS_SOURCE_ID:xlnneg006b06 5' similar to
            SW:PAR2_RAT Q63645 PROTEINASE ACTIVATED RECEPTOR 2 PRECURSOR ;,
            mRNA sequence.
ACCESSION   BF426487
VERSION     BF426487.1  GI:11437917
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
            , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
            ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Xenopus EST project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Sandy Clifton, Ph.D.
            WashU Xenopus EST project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by Bruce Blumberg
             Library normalized by Jihwan Song
             DNA Sequencing by: Washington University Genome Sequencing Center
             This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 296.
FEATURES             Location/Qualifiers
     source          1. .531
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XENOPUS_SOURCE_ID:xlnneg006b06"
```

```
                     /clone_lib="Xenopus laevis unfertilized egg cDNA library"
                     /tissue_type="unfertilized egg"
                     /lab_host="Top-10 F'"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; This library has been used successfully to clone a
                     number of full-length cDNAs ranging in size from 1.4 to
                     4.5 kb. There are less than 0.5% clones with multiple
                     inserts. Since each cDNA has an EcoRI site at its 5' end
                     and an XhoI site at the 3' end, these clones can be easily
                     identified. One should be suspicious of any clone which
                     gives 3 or more bands in an EcoRI-XhoI double digest AND
                     has an internal XhoI site. We usually do not further
                     characterize any such clones unless the cDNA is known to
                     give multiple bands in an EcoRI-XhoI digest. Microplate
                     status: 500,000 unamplified cDNAs were mass excised
                     (pBluescript SK-) in X11-Blue using ExAssist phage. The
                     resulting single-stranded phagemids were used to infect
                     Top10F'. Clones were picked into freezing medium (per
                     liter 15 g tryptone, 10g yeast extract, 5g NaCl, 36 mM
                     K2HPO4, 13.2 mM KH2PO4, 1.7 mM Na-citrate, 0.4 mM MgSO4*7
                     H2O, 6.8 mM (NH4)2SO4, 4% w/v glycerol) and grown for 24
                     hours. Original library contruction by Bruce Blumberg
                     (Blumberg et al., 1991 Science 253, 194-196; Hawlet et al.
                     , 1995, Genes Dev. 9, 2923-2935). Note: This is a Xenopus
                     Gene Collection (XGC) library"
BASE COUNT      133 a    121 c    102 g    175 t
ORIGIN

Alignment Scores:
Pred. No.:              3.1e-29       Length:         531
Score:                  369.50        Matches:        73
Percent Similarity:     61.54%        Conservative:   31
Best Local Similarity:  43.20%        Mismatches:     58
Query Match:            17.13%        Indels:         7
DB:                     12            Gaps:           5

US-09-208-629F-6 (1-408) x BF426487 (1-531)

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
              :::|||:::||||||:::   ||||||:::   |||||||||||:::
Db          1 TTACCTAGCAACGCCATTGCATTGTGGGTCTTTTTTTTCCGGACAAAGAAGAAACACCCT 60

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                  :::   |||      |||||||||:::||||||   :::|||   :::   ||||||
Db         61 GCCATGATTTATATGGCCAACTTGGCACTGGCAGATTTAATGTTTGTCATATGGCTCCCT 120

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||   ||||||||||||||||||||||||   :::||||||   ||||||:::
Db        121 TTTAAAATTGAATATCATCTGAATGGAAACAACTGGACTTATGGAGAACCTTTATGCAAA 180

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                          |||||||||||||||||||||||||||||   :::   |||:::
Db        181 GTTTTGATTGGGTTCTTCTATGGGAACATGTATTGCTCCATTCTTTTCATGACATGTCTT 240

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
              |||:::   ||||||      ||||||:::|||   ::::::              |||:::
Db        241 AGTGTCCAGAGGTACTGGGTCATCGTGAACCCCATGTCCCAC------ACAAGAAAGAAT 294

Qy        227 Thr---TyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
              |||      |||::::::              :::|||      :::             :::
Db        295 ACCAAACGTGCTATGATTGTTTCAATTGCTATATGGGTGGTCATTATGCATGGCACTATT 354

Qy        246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
              |||   :::::::::   |||      ||||||            ||||||||||||||
Db        355 CCATTGTACCTGATTAACCAAACTCTGTATCTGTCTGACCTTGGTATCACAACTTGCCAT 414

Qy        266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr---TyrPheIleSer 284
              ||||||                  ||:::          ||||||   |||:::::::::
Db        415 GACGTCTTGCCGCTG------GATTCTGCAACACTTGACCTGTATCACTACTACTTGGCT 468

Qy        285 LeuAlaPhePheGlyPheLeuIlePro 293
              ||||||      :::|||   |||
Db        469 CTAGCCCATTGAGTGTATCTTTCTCCA 495
```

```
RESULT 38
AF345566
LOCUS       AF345566               1199 bp    mRNA    linear   HTC 13-JUL-2001
DEFINITION  Homo sapiens putative G-protein-coupled receptor FKSG78 (FKSG78)
            mRNA, complete cds.
ACCESSION   AF345566
VERSION     AF345566.1  GI:13517959
KEYWORDS    HTC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1199)
  AUTHORS   Wang,Y.-g. and Gong,L.
  TITLE     Identification of FKSG78, a novel gene encoding a putative
            G-protein-coupled receptor
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1199)
  AUTHORS   Wang,Y.-g.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2001) Beijing FENGKESHENG Function Gene
            Technology Ltd., 4 Tou Tiao Lu Chang Street, Xuanwu District,
            Beijing 100050, P.R. China
FEATURES             Location/Qualifiers
     source          1. .1199
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /tissue_type="heart"
                     /dev_stage="fetus"
     gene            1. .1199
                     /gene="FKSG78"
     CDS             307. .1116
                     /gene="FKSG78"
                     /codon_start=1
                     /product="putative G-protein-coupled receptor FKSG78"
                     /protein_id="AAK29069.1"
                     /db_xref="GI:13517960"
                     /translation="MTLTFPFRIVHDAGFGPWYFKFILCRYTSVLFYANMYTSIVFLG
                     LISIDRYLKVVVKPFGDSRMYSITFTKVLSVCVWVIMAVLSLPNIILTNGQPTEDNIHD
                     CSKLKSPLGVKWHTAVTYVNSCLFVAVLVILIGCYIAISRYIHKSSRQFISQSSRKRK
                     HNQSIRVVVAVFFTCFLPYHLCRIPFTFSHLDRLLDESAQKILYYCKEITLFLSACNV
                     CLDPIIYFFMCRSFSRRLFKKSNIRTRSESIRSLQSVRRSEVRIYYDYTDV"
BASE COUNT      334 a    249 c    245 g    371 t
ORIGIN

Alignment Scores:
Pred. No.:              1.25e-28      Length:         1199
Score:                  368.50        Matches:        89
Percent Similarity:     50.33%        Conservative:   64
Best Local Similarity:  29.28%        Mismatches:     116
Query Match:            17.08%        Indels:         35
DB:                     11            Gaps:           9

US-09-208-629F-6 (1-408) x AF345566 (1-1199)

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
              ||||||   |||          |||   :::          |||   ::::::|||   :::
Db        133 AAGAACACCACC---------CTTCACAATGAATTTGACACAATTGTCTTGCCGGTGCTT 183

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              ||||||::::::||||||      :::          |||   :::   :::|||   :::|||
Db        184 TATCTCATTATATTTGTGGCAAGCATCTTGCTGAATGGTTTAGCAGTGTGG---ATCTTC 240

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||      |||:::      :::   :::||||||      |||:::   :::||||||
Db        241 TTCCACATTAGGAATAAAACCAGCTTCATATTCTATCTCAAAAACATAGTGGTTGCAGAC 300

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
                  :::      :::|||   ||||||:::|||   :::                  |||
Db        301 CTCATAATGACGCTGACATTTCCATTTCGAATAGTCCATGATGCAGGATTTGGACCTTGG 360
```

```
Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              |||     :::|||||||||   |||:::|||:::||||||  |||||||||
Db        361 TACTTCAAGTTTATTCTCTGCAGATACACTTCAGTTTTGTTTTATGCAAACATGTATACT 420

Qy        199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              ||||||:::   |||      |||||||||:::|||||||||   :::|||   ||||||
Db        421 TCCATCGTGTTCCTTGGGCTGATAAGCATTGATCGCTATCTGAAGGTGGTCAAGCCATTT 480

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
                      :::         |||:::      |||           ||||||
Db        481 GGGGACTCTCGGATGTACAGCATAACCTTCACGAAGGTTTTATCTGTTTGTGTTTGGGTG 540

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
              :::   :::      ||||||      ||||||                    |||
Db        541 ATCATGGCTGTTTTGTCTTTGCCAAACATCATCCTAACAAATGGT-----------CAG 588

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
              |||         |||      ::::::|||:::      :::      ||||||
Db        589 CCA---------ACAGAGGACAATATCCATGACTGCTCAAAACTTAAAAGTCCTTTGGGG 639

Qy        279 LeuTyrTyrPheIleSerLeuAlaPheGly---------PheLeuIleProPheVal 295
              :::   :::      ::::::   :::                |||:::         |||
Db        640 GTCAAATGGCATACGGCAGTCACCTATGTGAACAGCTGCTTGTTTGTGGCCGTGCTGGTG 699

Qy        296 LeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrp 315
              ::::::|||   ||||||   ||||||   |||   ::::::              ::::::
Db        700 ATTCTGATCGGATGTTACATAGCCATATCCAGGTACATCCACAAATCCAGCAGGCAATTC 759

Qy        316 Leu--------------------TrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
              :::                     :::   ::::::      :::::::::   :::
Db        760 ATAAGTCAGTCAAGCCGAAAGCGAAAACATAACCAGAGCATCAGGGTTGTTGTGGCTGTG 819

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
              |||      ||||||   |||   ::::::      |||                 :::
Db        820 TTTTTTACCTGCTTTCTACCATATCACTTGTGCAGAATT-----------CCTTTTACT 867

Qy        349 TyrAsnAsnThrAspGly-------------------------LeuTyrPheIleTyr 359
              :::::::::   |||                          ||||||:::
Db        868 TTTAGTCACTTAGACAGGCTTTTAGATGAATCTGCACAAAAAATCCTATATTACTGCAAA 927

Qy        360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379
              |||   |||   |||   :::   |||   ||||||||||||   :::||||||
Db        928 GAAATTACACTTTTCTTGTCTGCGTGTAATGTTTGCCTGGATCCAATAATTTACTTTTTC 987

Qy        380 MetSerLysThr 383
              |||   ::::::
Db        988 ATGTGTAGGTCA 999

RESULT 39
AL532537
LOCUS       AL532537                904 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL532537 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DM006YC06 5
            prime, mRNA sequence.
ACCESSION   AL532537
VERSION     AL532537.1  GI:12796030
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DM006YC06"
                     /clone_lib="LTI_NFL001_NBC4"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      260 a    220 c    160 g    264 t
ORIGIN

Alignment Scores:
Pred. No.:              2.55e-28        Length:         904
Score:                  364.00          Matches:        82
Percent Similarity:     49.81%          Conservative:   49
Best Local Similarity:  31.18%          Mismatches:     88
Query Match:            16.88%          Indels:         44
DB:                     9               Gaps:           7

US-09-208-629F-6 (1-408) x AL532537 (1-904)

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
              |||         |||   :::      ::::::|||      |||   |||||||||:::
Db        146 TATGCACATCACAGCACGGCCAGGATAGTAATGCCTCTGCATTACAGCCTCGTCTTCATC 205

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
              :::|||:::      |||   :::   |||   ::::::           :::   |||
Db        206 ATTGGGCTCGTGGGAAACTTACTAGCCTTGGTCGTCATTGTTCAAAACAGGAAAAAAATC 265

Qy        146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
                    ||||||::::::   |||||||||   |||:::|||   ||||||
Db        266 AACTCTACCACCCTCTATTCAACAAATTTGGTGATTTCTGATATACTTTTTACCACCGCT 325

Qy        165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
              ||||||   :::|||||||||:::      |||   :::|||      |||:::   |||
Db        326 TTGCCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGAGAATCGGAGATGCCTTG 385

Qy        185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAla 204
              ||||||   |||   :::::||||||   |||   |||      :::      :::
Db        386 TGTAGGATAACTGCGCTAGTGTTTTACATCAACACATATGCAGGTGTGAACTTTATGACC 445

Qy        205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224
              |||:::||||||:::|||::::::|||:::|||||||||      |||      :::
Db        446 TGCCTGAGTATTGACCGCTTCATTGCTGTGGTGCACCCTCTACGCTACAACAAGATAAAA 505

Qy        225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
              :::      :::|||         |||      ||||||      ||||||
Db        506 AGGATTGAACATGCAAAAGGCGTGTGCATATTTGTCTGGATTCTAGTATTTGCTCAGACA 565

Qy        245 LeuProPhePheIle---------LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
              ||||||      |||            |||||||||                    :::
Db        566 CTCCCACTCCTCATCAACCCTATGTCAAAGCAGGAG--------------GCTGAAAGG 610

Qy        262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer---ProPheGlnLeuTyr 280
              ||||||   :::      |||      |||:::   |||   |||:::   |||
Db        611 ATTACATGCATGGAGTATCCAAACTTTGAAGAAACTAAATCTCTTCCCTGGATTCTG--- 667

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
                    :::      |||   |||:::::::::|||   ::::::|||:::   |||
Db        668 ------CTTGGGGCATGTTTCATAGGATATGTACTTCCACTTATAATCATTCTCATCTGC 721

Qy        301 TyrAlaAlaIleIle------------------------------------------- 305
              |||:::   |||
Db        722 TATTCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAAACCCACTCACTGAGAAA 781

Qy        306 --------------ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
                           :::   ||||||
```

```
Db        782 TCTGGTGTAAACAAAAAGGCTCTCAACACAATT-------------------------- 814

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
                  :::||||||:::|||:::|||   :::||||||   |||   ::::::   :::
Db        815 ------ATTCTTATTATTGTTGTGTTTGTTCTCTGTTTCACACCTTACCATGTTGCAATT 868

Qy        341 IleIleHis 343
                  |||   |||
Db        869 ATTCAACAT 877

RESULT 40
AL547762
LOCUS       AL547762                931 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL547762 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI017YN05 5
            prime, mRNA sequence.
ACCESSION   AL547762
VERSION     AL547762.1  GI:12882129
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 931)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
      .     Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .931
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI017YN05"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      268 a    224 c    165 g    273 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.65e-28        Length:         931
Score:                  364.00          Matches:        82
Percent Similarity:     49.81%          Conservative:   49
Best Local Similarity:  31.18%          Mismatches:     88
Query Match:            16.88%          Indels:         44
DB:                     9               Gaps:           7

US-09-208-629F-6 (1-408) x AL547762 (1-931)

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
                  |||         |||   :::      :::::::|||       |||   |||||||||:::
Db        121 TATGCACATCACAGCACGGCCAGGATAGTAATGCCTCTGCATTACAGCCTCGTCTTCATC 180

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
                  :::|||:::      |||   :::   |||   ::::::             :::   |||
Db        181 ATTGGGCTCGTGGGAAACTTACTAGCCTTGGTCGTCATTGTTCAAAACAGGAAAAAAATC 240

Qy        146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
                     ||||||::::::   |||||||||   |||:::|||   ||||||
Db        241 AACTCTACCACCCTCTATTCAACAAATTTGGTGATTTCTGATATACTTTTTACCACCGCT 300

Qy        165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
                  ||||||   :::|||||||||:::      |||   :::|||      |||:::   |||
Db        301 TTGCCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGAGAATCGGAGATGCCTTG 360

Qy        185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAla 204
                  ||||||   |||   ::::::||||||   |||   |||       :::       :::
Db        361 TGTAGGATAACTGCGCTAGTGTTTTACATCAACACATATGCAGGTGTGAACTTTATGACC 420

Qy        205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224
                  |||:::|||||||:::|||:::::::|||:::|||||||||       |||       :::
Db        421 TGCCTGAGTATTGACCGCTTCATTGCTGTGGTGCACCCTCTACGCTACAACAAGATAAAA 480

Qy        225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
                  :::       :::|||         |||       ||||||       ||||||
Db        481 AGGATTGAACATGCAAAAGGCGTGTGCATATTTGTCTGGATTCTAGTATTTGCTCAGACA 540

Qy        245 LeuProPhePheIle---------LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
                  ||||||      |||                 |||||||||                 :::
Db        541 CTCCCACTCCTCATCAACCCTATGTCAAAGCAGGAG----------------GCTGAAAGG 585

Qy        262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer---ProPheGlnLeuTyr 280
                  ||||||   :::       |||       |||:::   |||   |||:::   |||
Db        586 ATTACATGCATGGAGTATCCAAACTTTGAAGAAACTAAATCTCTTCCCTGGATTCTG--- 642

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
                  :::          |||   |||:::::::::|||   :::::::|||:::   |||
Db        643 ------CTTGGGGCATGTTTCATAGGATATGTACTTCCACTTATAATCATTCTCATCTGC 696

Qy        301 TyrAlaAlaIleIle---------------------------------------------- 305
                  |||:::   |||
Db        697 TATTCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAAACCCACTCACTGAGAAA 756

Qy        306 ---------------ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
                              :::   ||||||
Db        757 TCTGGTGTAAACAAAAAGGCTCTCAACACAATT-------------------------- 789

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
                  :::||||||:::|||:::|||   :::||||||   |||   ::::::   :::
Db        790 ------ATTCTTATTATTGTTGTGTTTGTTCTCTGTTTCACACCTTACCATGTTGCAATT 843

Qy        341 IleIleHis 343
                  |||   |||
Db        844 ATTCAACAT 852


Search completed: June 29, 2003, 10:14:45
Job time : 1226.48 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 09:48:59 ; Search time 132.162 Seconds
(without alignments)
4582.672 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIKM.....AYLTKXXNDLREQGQPSQRT 408

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09208629/runat_24062003_101628_13027/app_query.fasta_1.1166
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09208629_@CGN_1_1_153_@runat_24062003_101628_13027
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2142	99.3	1830	9	US-10-225-567A-301	Sequence 301, App
2	586.5	27.2	1451	9	US-10-225-567A-299	Sequence 299, App
3	523.5	24.3	4895	9	US-10-225-567A-515	Sequence 515, App
4	523	24.2	3299	9	US-10-208-408-24	Sequence 24, Appl
5	523	24.2	3592	9	US-10-225-567A-323	Sequence 323, App
6	488	22.6	2588	10	US-09-943-718-3	Sequence 3, Appli
7	486.5	22.6	1080	10	US-09-943-718-5	Sequence 5, Appli
8	465	21.6	1425	9	US-10-094-417-9	Sequence 9, Appli
9	465	21.6	1955	9	US-10-190-469-2	Sequence 2, Appli
10	465	21.6	2137	9	US-09-782-974C-75	Sequence 75, Appl
11	465	21.6	2137	9	US-10-225-567A-529	Sequence 529, App
12	460	21.3	1080	10	US-09-739-151-1	Sequence 1, Appli
13	413.5	19.2	1020	9	US-10-251-385-181	Sequence 181, App
14	410.5	19.0	1020	9	US-10-251-385-31	Sequence 31, Appl
15	410.5	19.0	1020	10	US-09-788-133-1	Sequence 1, Appli
16	410.5	19.0	1086	9	US-10-222-024-1	Sequence 1, Appli
17	410.5	19.0	1086	9	US-10-251-385-77	Sequence 77, Appl
18	410.5	19.0	1086	9	US-10-251-385-205	Sequence 205, App
19	410.5	19.0	1638	9	US-10-225-567A-111	Sequence 111, App
20	410.5	19.0	2070	9	US-10-225-567A-303	Sequence 303, App
21	401.5	18.6	2424	9	US-10-225-567A-218	Sequence 218, App
22	389	18.0	1301	9	US-10-024-494-7	Sequence 7, Appli
c 23	387	17.9	1163	9	US-10-121-101B-9	Sequence 9, Appli
24	387	17.9	2299	9	US-10-225-567A-224	Sequence 224, App
25	386.5	17.9	1429	9	US-09-077-173A-1	Sequence 1, Appli
26	384	17.8	1118	9	US-10-225-567A-220	Sequence 220, App
27	382	17.7	1098	9	US-10-225-567A-331	Sequence 331, App
28	376.5	17.5	1041	9	US-09-828-478-1	Sequence 1, Appli
29	376.5	17.5	1041	10	US-09-826-791-5	Sequence 5, Appli
30	376.5	17.5	1041	10	US-09-866-230-6	Sequence 6, Appli
31	376.5	17.5	1260	9	US-09-779-679-1	Sequence 1, Appli
32	376.5	17.5	1260	9	US-09-779-679-24	Sequence 24, Appl
33	376.5	17.5	1430	9	US-09-828-478-3	Sequence 3, Appli
34	376.5	17.5	1700	10	US-09-728-952-26	Sequence 26, Appl
35	376.5	17.5	2807	9	US-10-225-567A-588	Sequence 588, App
36	374.5	17.4	993	10	US-09-826-791-1	Sequence 1, Appli
37	372.5	17.3	1014	9	US-10-225-567A-546	Sequence 546, App
38	371	17.2	1358	9	US-10-167-192-4	Sequence 4, Appli
39	368.5	17.1	1077	10	US-09-826-508-9	Sequence 9, Appli
40	368.5	17.1	1402	9	US-10-225-567A-417	Sequence 417, App
41	368.5	17.1	1481	9	US-10-071-766-136	Sequence 136, App
42	366	17.0	2025	9	US-10-225-567A-216	Sequence 216, App
43	365.5	16.9	1146	9	US-10-225-567A-413	Sequence 413, App
44	362	16.8	1697	9	US-10-109-533A-1	Sequence 1, Appli
45	361	16.7	1444	9	US-09-974-298-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-10-225-567A-301
; Sequence 301, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 301
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-301

Alignment Scores:

Pred. No.:	5.89e-189	Length:	1830
Score:	2142.00	Matches:	405
Percent Similarity:	99.26%	Conservative:	0
Best Local Similarity:	99.26%	Mismatches:	3

```
Query Match:            99.30%      Indels:        0
DB:                     9           Gaps:          0

US-09-208-629F-6 (1-408) x US-10-225-567A-301 (1-1830)

Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet  20
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db         88 TGCTCCATGATTTTACAGATTTCATAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 147

Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer  40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 207

Qy         41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 267

Qy         61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly  80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 327

Qy         81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 387

Qy        101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 447

Qy        121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTCTTC 507

Qy        141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 AGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTTCTT 567

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTATTT 627

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 687

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 CTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTAC 747

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACAGTT 807

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 867

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 ATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTCTAT 927

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTACTGC 987

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 1047

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 GCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATTCTT 1107

Qy        341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATATCTC 1167

Qy        361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 ATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTCATG 1227

Qy        381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
              ||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db       1228 TCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGAGAA 1287

Qy        401 GlnGlyGlnProSerGlnArgThr 408
              |||||||||||||||||||||||
Db       1288 CAAGGACAGCCATCACAGAGAACG 1311

RESULT 2
US-10-225-567A-299
; Sequence 299, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 299
;   LENGTH: 1451
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-299

Alignment Scores:
Pred. No.:              5.93e-45        Length:         1451
Score:                  586.50          Matches:        141
Percent Similarity:     50.61%          Conservative:   67
Best Local Similarity:  34.31%          Mismatches:     146
Query Match:            27.19%          Indels:         57
DB:                     9               Gaps:           14

US-09-208-629F-6 (1-408) x US-10-225-567A-299 (1-1451)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla  27
              |||   |||:::|||      :::          :::   |||   |||      |||
Db        139 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 198

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn  47
              |||   |||                    |||   ||||||          ||||||
Db        199 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 237

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe  67
                 :::|||            |||              |||
Db        238 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA-------------- 282

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle  87
                       |||   :::   |||      |||   |||:::
Db        283 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 324

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                     |||   |||||||||                                 |||
Db        325 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 354

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
              |||      |||:::|||      :::|||   :::|||   :::|||||||||||||||
Db        355 ACTGGAAAACTGACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGT 414
```

```
Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
              :::|||:::|||   :::   ||||||:::      |||||||||:::
Db        415 TTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 474

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
              |||   |||      |||||||||:::||||||   |||      :::      |||
Db        475 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 534

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              ||||||||||||||||:::::   |||||||||:::::||||||   ||||||
Db        535 TTGAAGATTGCCTATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 594

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
              |||||||||||||||||||||||||||||||   :::   |||:::
Db        595 GTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 654

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
              |||:::   ||||||      ||||||:::|||      :::            |||
Db        655 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 711

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
              |||:::                  :::|||      :::   |||      :::|||
Db        712 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCT 771

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
              :::::::::||||||      ::::::      :::||||||||||||||||||||
Db        772 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 831

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr----------------- 280
              |||                     |||   ||||||
Db        832 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 870

Qy        281 TyrPheIleSerLeuAlaPheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||:::|||||||||      ||||||   |||      |||
Db        871 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 930

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla----------------TyrAspHisArg 314
              |||      :::||||||   |||   :::                  :::   :::
Db        931 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 990

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
                  :::|||      :::   :::|||   ::::::   |||||||||
Db        991 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACT 1050

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
              |||||||||:::::|||:::::|||      |||:::      :::   |||
Db       1051 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1101

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
                     :::|||   :::|||:::::|||||||||||||   :::||||||||||||
Db       1102 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1161

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
              :::|||||||||:::|||:::   :::|||         |||:::|||:::      |||
Db       1162 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1221

Qy        390 TyrLeuThrLys******AsnAspLeuArgGlu 400
              |||   :::          :::::::::
Db       1222 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1254

RESULT 3
US-10-225-567A-515
; Sequence 515, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 515
;   LENGTH: 4895
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-515

Alignment Scores:
Pred. No.:                2.26e-38        Length:         4895
Score:                    523.50          Matches:        121
Percent Similarity:       53.57%          Conservative:   59
Best Local Similarity:    36.01%          Mismatches:     145
Query Match:              24.27%          Indels:         12
DB:                       9               Gaps:           5

US-09-208-629F-6 (1-408) x US-10-225-567A-515 (1-4895)

Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
              ||||||   |||      |||      |||   |||   |||||||||
Db        225 CTGGCGGCACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGA-GGTGGTGAT 283

Qy         68 GluGluPhePro------PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85
              :::      |||      |||      |||:::   |||
Db        284 GACAGCACGCCCTCAATCCTGCCTGCCCCCCGCGGCTACCCAGGCCAAGTC--------- 334

Qy         86 LysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105
                        |||      :::|||   :::   |||   :::   :::::::::
Db        335 ---------TGTGCCAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGGCA 382

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
                  |||          :::   |||:::|||:::||||||:::|||   ||||||   |||
Db        383 CTGCTTCTGGGCTGGGTGCCCACCAGGCTGGTGCCCGCCCTCTATGGGCTGGTCCTGGTG 442

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
              ||||||:::||||||||||   :::   ||||||:::|||      :::      :::
Db        443 GTGGGGCTGCCGGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACGCAGGCACCTCGGCTG 502

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
                 :::|||:::          |||||||||   ||||||   |||      :::   |||
Db        503 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCCTGCTGGCCCTGGCGCTG 562

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
              |||   :::|||||||||||||||   |||      |||   |||||||||      |||
Db        563 CCCCCGCGGATCGCCTACCACCTGCGTGGCCAGCGCTGGCCCTTCGGGGAGGCCGCCTGC 622

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205
              |||      |||          ||||||:::||||||   |||:::||||||||||||
Db        623 CGCCTGGCCACGGCCGCACTCTATGGTCACATGTATGGCTCAGTGCTGCTGCTGGCCGCC 682

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
              :::|||::::::||||||||||||:::|||||||||          |||   |||
Db        683 GTCAGCCTGGATCGCTACCTGGCCCTGGTGCACCCGCTGCGGGCCCGCGCCCTGCGTGGC 742

Qy        226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
              |||||||      |||      |||                              |||
Db        743 CGGCGCCTGGCCCTTGGACTCTGCATGGCTGCTTGGCTCATGGCGGCCGCCCTGGCACTG 802

Qy        246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
              |||      :::   :::|||   :::   |||   :::   |||          ||||||
Db        803 CCCCTGACACTGCAGCGGCAGACCTTCCGGCTGGCGCGCTCCGATCGCGTGCTCTGCCAT 862

Qy        266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
              |||                     :::|||   :::|||          |||      |||
Db        863 GACGCGCTGCCCCTGGACGCACAGGCCTCCCACTGGCAA---CCGGCCTTCACCTGCCTG 919

Qy        286 AlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305
              |||      |||      :::|||   :::   ::::::   ||||||   |||   :::
Db        920 GCGCTGTTGGGCTGTTTCCTGCCCCTGCTGGCCATGCTGCTGTGCTACGGGGCCACCCTG 979

Qy        306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325
```

```
Db        980 CACACGCTGGCGGCCAGCGGCCGGCGCTACGGCCACGCGCTGAGGCTGACCGCAGTGGTG 1039

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345

Db       1040 CTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGCAACCTGCTGCTGCTGCTGCATTACTCG 1099

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365

Db       1100 GACCCGAGCCCCAGCGCCTGGGGCAACCTCTATGGTGCCTACGTGCCCAGCCTGGCGCTG 1159

Qy        366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381

Db       1160 AGCACCCTCAACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1207

RESULT 4
US-10-208-408-24
; Sequence 24, Application US/10208408
; Publication No. US20030096272A1
; GENERAL INFORMATION:
;  APPLICANT: Schebye, Xiao Min
;  TITLE OF INVENTION: GENES REGULATED BY PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR
;  FILE REFERENCE: PA-0048-1 US
;  CURRENT APPLICATION NUMBER: US/10/208,408
;  CURRENT FILING DATE:  2002-07-29
;  PRIOR APPLICATION NUMBER: 60/308,868
;  PRIOR FILING DATE: 2001-07-30
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PERL Program
; SEQ ID NO 24
;   LENGTH: 3299
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030096272A1 2456481CB1
US-10-208-408-24

Alignment Scores:
Pred. No.:              1.44e-38        Length:         3299
Score:                  523.00          Matches:        129
Percent Similarity:     52.58%          Conservative:   75
Best Local Similarity:  33.25%          Mismatches:     150
Query Match:            24.25%          Indels:         34
DB:                     9               Gaps:           12

US-09-208-629F-6 (1-408) x US-10-208-408-24 (1-3299)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33

Db         62 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 121

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53

Db        122 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 169

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72

Db        170 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 229

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----------IleThrValLysIleLys 88

Db        230 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 289

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108

Db        290 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 349

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128

Db        350 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 409

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146

Db        410 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 466

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166

Db        467 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 526

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186

Db        527 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 586

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206

Db        587 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 646

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223

Db        647 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 706

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243

Db        707 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 757

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263

Db        758 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 817

Qy        264 CysHisAspValHisAsn------ThrCysGluSerSerSerProPheGlnLeuTyrTyr 281

Db        818 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGCTACTATCCC---------TACTAC 868

Qy        282 PheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyr 301

Db        869 TTCTCAGCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTAT 928

Qy        302 AlaAlaIleIleArgThrLeuAsn--------------------AlaTyrAspHisArg 314

Db        929 GTGTCTATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGG 988

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334

Db        989 GCTTTGTTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGA 1039

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAsp 353

Db       1040 CCCACAAACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAG 1099

Qy        354 GlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373

Db       1100 GCTGCCTACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGAC 1159

Qy        374 ProPheLeuTyrPheLeuMetSer 381

Db       1160 CCCCTAATTTACTATTACGCTTCC 1183

RESULT 5
US-10-225-567A-323
; Sequence 323, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 323
```

```
;   LENGTH: 3592
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-323

Alignment Scores:
Pred. No.:              1.63e-38        Length:         3592
Score:                  523.00          Matches:        127
Percent Similarity:     52.59%          Conservative:   76
Best Local Similarity:  32.90%          Mismatches:     153
Query Match:            24.25%          Indels:         30
DB:                     9               Gaps:           11

US-09-208-629F-6 (1-408) x US-10-225-567A-323 (1-3592)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||       |||:::       |||   ||||||
Db        348 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 407

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||     :::                     ||||||                |||
Db        408 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 455

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||     ::::::|||             ||||||     :::|||   |||
Db        456 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 515

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr----------IleThrValLysIleLys 88
                 :::     :::|||   |||             :::::::::
Db        516 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 575

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
              |||   :::                 :::                ||||||||||||
Db        576 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 635

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||      :::|||::::::|||      ||||||||||   :::
Db        636 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 695

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||      |||:::   :::::::::|||      :::   :::
Db        696 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 752

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::      :::||||||   ||||||   ||||||          ||||||
Db        753 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 812

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||::::::|||   ||||||      ||||||||
Db        813 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 872

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                 |||     ||||||   |||||||||   ||||||||||:::      |||
Db        873 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 932

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::|||||||:::|||:::|||           ::::::|||   |||
Db        933 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 992

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
                 :::   ::::::          ||||||      :::||||||
Db        993 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 1043

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              :::::::|||      :::   :::|||         :::         :::||||||||||
Db       1044 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 1103

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                     :::       |||||||||
Db       1104 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1160

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::         |||   :::|||   ::::::         ||||||   :::

Db       1161 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1220

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              |||||||||   |||:::                   :::       |||   |||
Db       1221 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1280

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::           |||      :::      ||||||   |||||||||   |||:::
Db       1281 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1331

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||:::::||||||   |||::::::         :::   :::   |||:::
Db       1332 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1391

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||:::::||||||:::||||||
Db       1392 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1451

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1452 ATTTACTATTACGCTTCC 1469

RESULT 6
US-09-943-718-3
; Sequence 3, Application US/09943718
; Patent No. US20020103361A1
;   GENERAL INFORMATION:
;        APPLICANT: Huffine, Constance F.
;                   Rossi, Devora L.
;                   Capone, Myriam
;                   Hedrick, Joseph A.
;                   Vicari, Alain
;                   Gorman, Daniel M.
;                   Zlotnik, Albert
;        TITLE OF INVENTION: Mammalian Chemokines; Receptors;
;                            Reagents; Uses
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DNAX Research Institute
;             STREET: 901 California Avenue
;             CITY: Palo Alto
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94304-1104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/943,718
;             FILING DATE: 30-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/009,817
;             FILING DATE: 20-APR-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ching, Edwin P.
;             REGISTRATION NUMBER: 34,090
;             REFERENCE/DOCKET NUMBER: DX0588K
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (650)852-9196
;             TELEFAX: (650)496-1200
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2588 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..1080
```

```
;       FEATURE:
;            NAME/KEY:  misc_feature
;            LOCATION:  158
;            OTHER INFORMATION:  /note= "residues 158, 159, and 276
;            probably absent, changing reading frame between those positions;
;            sequences provided in SEQ ID NO: 5 and 6"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-943-718-3

Alignment Scores:
Pred. No.:              1.8e-35       Length:         2588
Score:                  488.00        Matches:        117
Percent Similarity:     51.45%        Conservative:   60
Best Local Similarity:  34.01%        Mismatches:     134
Query Match:            22.62%        Indels:         34
DB:                     10            Gaps:           4

US-09-208-629F-6 (1-408) x US-09-943-718-3 (1-2588)

Qy         56 IleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeu 75
              :::|||:::   ||||||:::
Db       1135 GTAAAGAGTACCAGGGGAAGCCAT---------------------------------- 1158

Qy         76 GluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer------ 93
              ||||||      |||         ||||||:::::|||   |||:::   |||
Db       1159 GAAGGCCCTCTGGGTCCC------ACAGTACAACTCAAGGAGCCGAAGTCCTCAGACAAG 1212

Qy         94 ----------------------------------------AlaSerHisLeuHisVal 99
                                                      :::      |||   :::
Db       1213 CCTAATCCACGAGGCTACCCGGGCAAATTCTGTGCCAACGACAGTGACACGCTGGAGCTC 1272

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
              ::::::      |||       :::   |||||||||:::||||||:::
Db       1273 CCGGCCAGCTCTCAAGCACTGCTGCTGGGGTGGGTCCCCACGAAGCTGGTACCTGCCCTC 1332

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   ||||||      ||||||:::|||||||||   :::   ||||||:::|||
Db       1333 TATGGGCTTGTGGTGGCTGTGGGGCTGCCTGCCAATGGGCTGGCGCTGTGGGTGCTGGCC 1392

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyrThr-AsnLeuAlaIleAlaAspPh 159
              |||         :::   :::|||:::          |||||||||:::||||||
Db       1393 ACAAGGGTGCCACGCCTGCCATCCACCATTCTGCTCATGGAACCTGGCAGTGGCTGATCT 1452

Qy        159 eLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVa 179
              |||      :::   ||||||   ::::::|||||||||||||   |||      |||
Db       1453 GCTGTTGGCCCTGGTGCTGCCACCACGACTGGCTTACCACTTGCGTGGCCAGCGCTGGCC 1512

Qy        179 lPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSe 199
              |||||||||      ||||||      |||         ||||||:::||||||   ||
Db       1513 ATTTGGTGAGGCTGCCTGCCGGGTGGCCACAGCTGCCCTCTATGGCCACATGTATGGTTC 1572

Qy        199 rIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheTh 219
              |:::||||||||||||   :::|||::::::||||||||||:::||||||||
Db       1573 AGTGTTGCTGCTGGCTGCAGTCAGCTTGGACAGATACCTGGCCCTGGTGCATCCTTTGCG 1632

Qy        219 rTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaTh 239
              |||   |||                     |||   :::   |||   ::
Db       1633 GGCCCGTGCACTGCGTGGTCAACGCCTCACTACTGGACTCTGTTTGGTGGCCTGGCTCTC 1692

Qy        239 rValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPr 259
              :               ||||||      :::   :::|||   :::   |||:::   ||
Db       1693 TGCAGCCACCCTGGCCTTGCCTCTCACTCTGCATCGGCAGAACTTCCGATTACTGGCTCC 1752

Qy        259 oAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLe 279
              |              |||||||||                |||   :::|||   ::::::
Db       1753 GAT-CGCATGCTGTGTCATGATGCGCTGCCCCTGGCTGAGCAGAACTCCCACTGGAGAAC 1811

Qy        279 uTyrTyrPheIleSerLeuAlaPhePheGlyPheLeu-IleProPheValLeuIleIleT 299
                  ||||||   ||||||      |||      :::|||   :::   :::
Db       1812 G---GCCTTCATCTGCCTGGCTGTCCTGGGCTGCTTCCTTGCCACTGCTGGCCATGGGCC 1868

Qy        299 yrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrV 319
                ||||||         :::|||   |||   |||         |||:::         :
```

```
Db       1869 TGTGCTATGGAACCACCCTTCGTGCATTGGCGGCCAATGGCCAGCGCTACAGCCATGCAC 1928

Qy        319 alLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleI 339
              :::::   :::   |||:::|||                |||   |||||||||::::
Db       1929 TCAGACTGACAGCCCTGGTACTGTTCTCGGCAGTGGCTTCTTTCACACCTAGCAATGTGC 1988

Qy        339 leLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleT 359
              ::|||::::::|||::::::|||                              ||||||      |
Db       1989 TGCTGGTGCTGCACTATTCAAACCCGAGCCCTGAGGCCTGGGGCAATCTCTATGGAGCCT 2048

Qy        359 yrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheL 379
              ||:::   :::|||   |||   :::|||||||||||:::|||||||||:::|||:::
Db       2049 ATGTGCCCAGCCTGGCACTCAGCACCCTCAACAGCTGCGTAGACCCTTTCATCTACTACT 2108

Qy        379 euMetSer 381
                :::|||
Db       2109 ATGTGTCC 2116

RESULT 7
US-09-943-718-5
; Sequence 5, Application US/09943718
; Patent No. US20020103361A1
;   GENERAL INFORMATION:
;        APPLICANT: Huffine, Constance F.
;                   Rossi, Devora L.
;                   Capone, Myriam
;                   Hedrick, Joseph A.
;                   Vicari, Alain
;                   Gorman, Daniel M.
;                   Zlotnik, Albert
;        TITLE OF INVENTION: Mammalian Chemokines; Receptors;
;                            Reagents; Uses
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DNAX Research Institute
;             STREET: 901 California Avenue
;             CITY: Palo Alto
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94304-1104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/943,718
;             FILING DATE: 30-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/009,817
;             FILING DATE: 20-APR-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ching, Edwin P.
;             REGISTRATION NUMBER: 34,090
;             REFERENCE/DOCKET NUMBER: DX0588K
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (650)852-9196
;             TELEFAX: (650)496-1200
;   INFORMATION FOR SEQ ID NO: 5:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1080 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..1077
;        SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-943-718-5

Alignment Scores:
```

Pred. No.:	7.23e-36	Length:	1080
Score:	486.50	Matches:	111
Percent Similarity:	52.96%	Conservative:	59
Best Local Similarity:	34.58%	Mismatches:	135
Query Match:	22.55%	Indels:	17
DB:	10	Gaps:	2

US-09-208-629F-6 (1-408) x US-09-943-718-5 (1-1080)

```
Qy         76 GluGlyTrpThrGlyAlaThrIleThrValLys-IleLysCysProGluGluSerAlaSe 95
Db          3 GAAGGCCCTCTGGGTCCCACAGTACAACTCAAGGAGCCGAAGTCCTCAGACAAGCCTAAT 62

Qy         95 rHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLys-- 114
Db         63 CCACGAGGCTACCCGGGCAAATTCTGTGCCAACGACAGTGACACGCTGGAGCTCCCGGCC 122

Qy        115 ----------------------------------LeuIleProAlaIleTyrLeuLe 122
Db        123 AGCTCTCAAGCACTGCTGCTGGGGTGGGTCCCCACAGCTGGTACCTGCCCTCTATGGGCT 182

Qy        122 uValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgTh 142
Db        183 TGTGGTGGCTGTGGGGCTGCCTGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACAAGGGT 242

Qy        142 rArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCy 162
Db        243 GCCACGCCTGCCATCCACCATTCTGCTCATGAACCTGGCAGTGGCTGATCTGCTGTTGGC 302

Qy        162 sValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGl 182
Db        303 CCTGGTGCTGCCACCACGACTGGCTTACCACTTGCGTGGCCAGCGCTGGCCATTTGGTGA 362

Qy        182 uValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLe 202
Db        363 GGCTGCCTGCCGGGTGGCCACAGCTGCCCTCTATGGCCACATGTATGGTTCAGTGTTGCT 422

Qy        202 uLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGl 222
Db        423 GCTGGCTGCAGTCAGCTTGGACAGATACCTGGCCCTGGTGCATCCTTTGCGGGCCCGTGC 482

Qy        222 yLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLe 242
Db        483 ACTGCGTGGTCAACGCCTCACTACTGGACTCTGTTTGGTGGCCTGGCTCTCTGCAGCCAC 542

Qy        242 uTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleTh 262
Db        543 CCTGGCCTTGCCTCTCACTCTGCATCGGCAGAACTTCCGATTACTGGCTCCGAT-CGCAT 601

Qy        262 rThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPh 282
Db        602 GCTGTGTCATGATGCGCTGCCCCTGGCTGAGCAGAACTCCCACTGGAGAACG---GCCTT 658

Qy        282 eIleSerLeuAlaPhePheGlyPheLeu-IleProPheValLeuIleIleTyrCysTyrA 302
Db        659 CATCTGCCTGGCTGTCCTGGGCTGCTTCCTTGCCACTGCTGGCCATGGGCCTGTGCTATG 718

Qy        302 laAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaS 322
Db        719 GAACCACCCTTCGTGCATTGGCGGCCAATGGCCAGCGCTACAGCCATGCACTCAGACTGA 778

Qy        322 erLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleI 342
Db        779 CAGCCCTGGTACTGTTCTCGGCAGTGGCTTCTTTCACACCTAGCAATGTGCTGCTGGTGC 838

Qy        342 leHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleA 362
Db        839 TGCACTATTCAAACCCGAGCCCTGAGGCCTGGGGCAATCTCTATGGAGCCTATGTGCCCA 898

Qy        362 laLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db        899 GCCTGGCACTCAGCACCCTCAACAGCTGCGTAGACCCTTTCATCTACTACTATGTGTCC 957
```

```
RESULT 8
US-10-094-417-9
; Sequence 9, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
;  APPLICANT: Tian, Hui
;  APPLICANT:  Zhao, Jiagang
;  APPLICANT:  Chen, Jin-Long
;  APPLICANT:  Cutler, Gene
;  APPLICANT:  Tularik Inc.
;  TITLE OF INVENTION: No. US20030045685A1el Receptors
;  FILE REFERENCE: 018781-008110US
;  CURRENT APPLICATION NUMBER: US/10/094,417
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: US 09/802,803
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: US 60/276,649
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;  LENGTH: 1425
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR341
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (241)..(1320)
;  OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR341
US-10-094-417-9
```

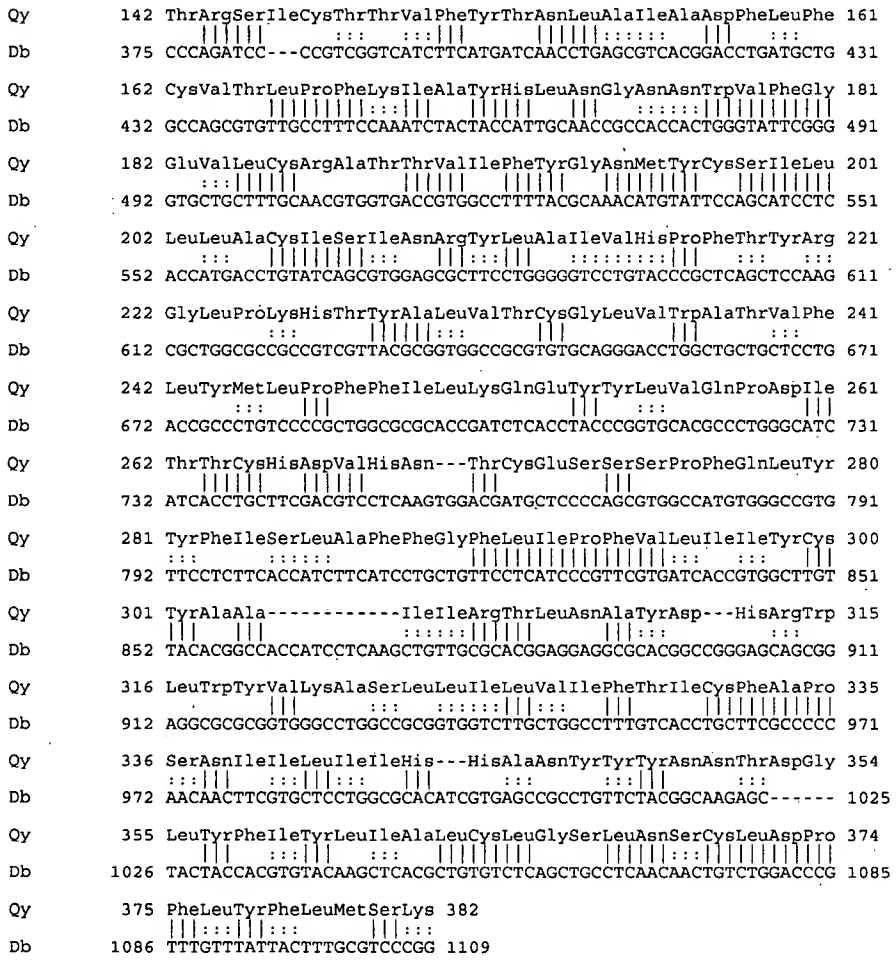
Pred. No.:	1.05e-33	Length:	1425
Score:	465.00	Matches:	104
Percent Similarity:	51.62%	Conservative:	55
Best Local Similarity:	33.77%	Mismatches:	131
Query Match:	21.56%	Indels:	18
DB:	9	Gaps:	8

US-09-208-629F-6 (1-408) x US-10-094-417-9 (1-1425)

```
Qy         90 ProGluGluSerAlaSerHisLeuHisVal-----------------LysAsnAlaThr 103
Db        220 CCGGGCGAGGACCCCTCCAGGATGCAGGTCCCGAACAGCACCGGCCCGGACAACGCGACG 279

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
Db        280 CTGCAGATGCTGCGGAACCCGGCGATCGCGGTGGCCCTGCCCGTGGTGTACTCGCTGGTG 339

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg------ 141
Db        340 GCGGCGGTCAGCATCCCGGGCAACCTCTTCTCTCTGTGGGTGCTGTGCCGGCGCATGGGG 399

Qy        142 ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPhe 161
Db        400 CCCAGATCC---CCGTCGGTCATCTTCATGATCAACCTGAGCGTCACGGACCTGATGCTG 456

Qy        162 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 181
Db        457 GCCAGCGTGTTGCCTTTCCAAATCTACTACCATTGCAACCGCCACCACTGGGTATTCGGG 516

Qy        182 GluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeu 201
Db        517 GTGCTGCTTTGCAACGTGGTGACCGTGGCCTTTTACGCAAACATGTATTCCAGCATCCTC 576

Qy        202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
Db        577 ACCATGACCTGTATCAGCGTGGAGCGCTTCCTGGGGGTCCTGTACCCGCTCAGCTCCAAG 636

Qy        222 GlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
Db        637 CGCTGGCGCCGCCGTCGTTACGCGGTGGCCGCGTGTGCAGGGACCTGGCTGCTGCTCCTG 696
```

```
Qy        242 LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db        697 ACCGCCCTGTCCCCGCTGGCGCGCACCGATCTCACCTACCCGGTGCACGCCCTGGGCATC 756

Qy        262 ThrThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        757 ATCACCTGCTTCGACGTCCTCAAGTGGACGATGCTCCCCAGCGTGGCCATGTGGGCCGTG 816

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db        817 TTCCTCTTCACCATCTTCATCCTGCTGTTCCTCATCCCGTTCGTGATCACCGTGGCTTGT 876

Qy        301 TyrAlaAla------------IleIleArgThrLeuAsnAlaTyrAsp---HisArgTrp 315
Db        877 TACACGGCCACCATCCTCAAGCTGTTGCGCACGGAGGAGGCGCACGGCCGGGAGCAGCGG 936

Qy        316 LeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 335
Db        937 AGGCGCGCGGTGGGCCTGGCCGCGGTGGTCTTGCTGGCCTTTGTCACCTGCTTCGCCCCC 996

Qy        336 SerAsnIleIleLeuIleIleHis---HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db        997 AACAACTTCGTGCTCCTGGCGCACATCGTGAGCCGCCTGTTCTACGGCAAGAGC------ 1050

Qy        355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
Db       1051 TACTACCACGTGTACAAGCTCACGCTGTGTCTCAGCTGCCTCAACAACTGTCTGGACCCG 1110

Qy        375 PheLeuTyrPheLeuMetSerLys 382
Db       1111 TTTGTTTATTACTTTGCGTCCCGG 1134

RESULT 9
US-10-190-469-2
; Sequence 2, Application US/10190469
; Publication No. US20030013155A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  TITLE OF INVENTION: 14400 Receptor, A No. US20030013155A1el G-Protein Coupled Recepto
;  FILE REFERENCE: 5800-7, 035800/169196
;  CURRENT APPLICATION NUMBER: US/10/190,469
;  CURRENT FILING DATE:  2002-07-05
;  PRIOR APPLICATION NUMBER: US/09/137,063A
;  PRIOR FILING DATE: 1999-08-20
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1955
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1955)
;   OTHER INFORMATION: n = a, t, c or g
US-10-190-469-2

Alignment Scores:
Pred. No.:              1.64e-33        Length:         1955
Score:                  465.00          Matches:        104
Percent Similarity:     51.62%          Conservative:   55
Best Local Similarity:  33.77%          Mismatches:     131
Query Match:            21.56%          Indels:         18
DB:                     9               Gaps:           8

US-09-208-629F-6 (1-408) x US-10-190-469-2 (1-1955)

Qy         90 ProGluGluSerAlaSerHisLeuHisVal-----------------LysAsnAlaThr 103
Db        267 CCGGGCGAGGACCCCTCCAGGATGCAGGTCCCGAACAGCACCGGCCCGGACAACGCGACG 326

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
Db        327 CTGCAGATGCTGCGGAACCCGGCGATCGCGGTGGCCCTGCCCGTGGTGTACTCGCTGGTG 386

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg------ 141
Db        387 GCGGCGGTCAGCATCCCGGGCAACCTCTTCTCTCTGTGGGTGCTGTGCCGGCGCATGGGG 446

Qy        142 ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPhe 161
Db        447 CCCAGATCC---CCGTCGGTCATCTTCATGATCAACCTGAGCGTCACGGACCTGATGCTG 503

Qy        162 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 181
Db        504 GCCAGCGTGTTGCCTTTCCAAATCTACTACCATTGCAACCGCCACCACTGGGTATTCGGG 563

Qy        182 GluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeu 201
Db        564 GTGCTGCTTTGCAACGTGGTGACCGTGGCCTTTTACGCAAACATGTATTCCAGCATCCTC 623

Qy        202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
Db        624 ACCATGACCTGTATCAGCGTGGAGCGCTTCCTGGGGGTCCTGTACCCGCTCAGCTCCAAG 683

Qy        222 GlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
Db        684 CGCTGGCGCCGCCGTCGTTACGCGGTGGCCGCGTGTGCAGGGACCTGGCTGCTGCTCCTG 743

Qy        242 LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db        744 ACCGCCCTGTCCCCGCTGGCGCGCACCGATCTCACCTACCCGGTGCACGCCCTGGGCATC 803

Qy        262 ThrThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        804 ATCACCTGCTTCGACGTCCTCAAGTGGACGATGCTCCCCAGCGTGGCCATGTGGGCCGTG 863

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db        864 TTCCTCTTCACCATCTTCATCCTGCTGTTCCTCATCCCGTTCGTGATCACCGTGGCTTGT 923

Qy        301 TyrAlaAla------------IleIleArgThrLeuAsnAlaTyrAsp---HisArgTrp 315
Db        924 TACACGGCCACCATCCTCAAGCTGTTGCGCACGGAGGAGGCGCACGGCCGGGAGCAGCGG 983

Qy        316 LeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 335
Db        984 AGCGCCGCGGTGGGCCTGGCCGCGGTGGTCTTGCTGGCCTTTGTCACCTGCTTCGCCCCC 1043

Qy        336 SerAsnIleIleLeuIleIleHis---HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db       1044 AACAACTTCGTGCTCCTGGCGCACATCGTGAGCCGCCTGTTCTACGGCAAGAGC------ 1097

Qy        355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
Db       1098 TACTACCACGTGTACAAGCTCACGCTGTGTCTCAGCTGCCTCAACAACTGTCTGGACCCG 1157

Qy        375 PheLeuTyrPheLeuMetSerLys 382
Db       1158 TTTGTTTATTACTTTGCGTCCCGG 1181

RESULT 10
US-09-782-974C-75
; Sequence 75, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
;  APPLICANT: Vogeli, Gabriel
;  APPLICANT:  Lind, Peter
;  APPLICANT:  Wood, Linda S.
;  APPLICANT:  Parodi, Luis A.
;  TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
;  FILE REFERENCE: 411USPHRM311
;  CURRENT APPLICATION NUMBER: US/09/782,974C
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: 60/165,838
;  PRIOR FILING DATE: 1999-11-16
```

```
;  PRIOR APPLICATION NUMBER: 09/714,449
;  PRIOR FILING DATE: 2000-11-16
;  PRIOR APPLICATION NUMBER: 60/198,568
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: 60/166,071
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/166,678
;  PRIOR FILING DATE: 1999-11-19
;  PRIOR APPLICATION NUMBER: 60/173,396
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/184,129
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/185,421
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/185,554
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/186,530
;  PRIOR FILING DATE: 2000-03-02
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 192
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
;   LENGTH: 2137
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-782-974C-75

Alignment Scores:
Pred. No.:                1.86e-33        Length:          2137
Score:                    465.00          Matches:         104
Percent Similarity:       51.62%          Conservative:    55
Best Local Similarity:    33.77%          Mismatches:      131
Query Match:              21.56%          Indels:          18
DB:                       9               Gaps:            8

US-09-208-629F-6 (1-408) x US-09-782-974C-75 (1-2137)

Qy          90 ProGluGluSerAlaSerHisLeuHisVal-----------------LysAsnAlaThr 103
               |||   |||      |||   :::   |||                 |||||||||
Db         195 CCGGGCGAGGACCCCTCCAGGATGCAGGTCCCGAACAGCACCGGCCCGGACAACGCGACG 254

Qy         104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
               :::      |||   :::                 :::|||   :::|||   ||||||
Db         255 CTGCAGATGCTGCGGAACCCGGCGATCGCGGTGGCCCTGCCCGTGGTGTACTCGCTGGTG 314

Qy         124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg------ 141
               |||   :::|||   |||      :::||||||:::|||      |||
Db         315 GCGGCGGTCAGCATCCCGGGCAACCTCTTCTCTCTGTGGGTGCTGTGCCGGCGCATGGGG 374

Qy         142 ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPhe 161
               ||||||      :::   :::|||      ||||||::::::   |||   :::
Db         375 CCCAGATCC---CCGTCGGTCATCTTCATGATCAACCTGAGCGTCACGGACCTGATGCTG 431

Qy         162 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 181
               |||||||||:::|||   ||||||   |||   ::::::||||||||||||
Db         432 GCCAGCGTGTTGCCTTTCCAAATCTACTACCATTGCAACCGCCACCACTGGGTATTCGGG 491

Qy         182 GluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeu 201
               :::||||||         ||||||   ||||||   |||||||||   |||||||||
Db         492 GTGCTGCTTTGCAACGTGGTGACCGTGGCCTTTTACGCAAACATGTATTCCAGCATCCTC 551

Qy         202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
               :::   |||||||||:::   |||:::|||   :::::::::|||   :::   :::
Db         552 ACCATGACCTGTATCAGCGTGGAGCGCTTCCTGGGGGTCCTGTACCCGCTCAGCTCCAAG 611

Qy         222 GlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
               :::      ||||||:::      |||         |||      :::
Db         612 CGCTGGCGCCGCCGTCGTTACGCGGTGGCCGCGTGTGCAGGGACCTGGCTGCTGCTCCTG 671

Qy         242 LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
               :::   |||                     |||   :::             |||
Db         672 ACCGCCCTGTCCCCGCTGGCGCGCACCGATCTCACCTACCCGGTGCACGCCCTGGGCATC 731

Qy         262 ThrThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPheGlnLeuTyr 280
               ||||||   ||||||         |||         |||
Db         732 ATCACCTGCTTCGACGTCCTCAAGTGGACGATGCTCCCCAGCGTGGCCATGTGGGCCGTG 791

Qy         281 TyrPheIleSerLeuAlaPheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
               :::      ::::::             ||||||||||||||||||:::   :::   |||
Db         792 TTCCTCTTCACCATCTTCATCCTGCTGTTCCTCATCCCGTTCGTGATCACCGTGGCTTGT 851

Qy         301 TyrAlaAla----------IleIleArgThrLeuAsnAlaTyrAsp---HisArgTrp 315
               |||   |||          ::::::|||||      |||:::         :::
Db         852 TACACGGCCACCATCCTCAAGCTGTTGCGCACGGAGGAGGCGCACGGCCGGGAGCAGCGG 911

Qy         316 LeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 335
                   |||      :::   ::::::|||:::   |||      ||||||||||||
Db         912 AGGCGCGCGGTGGGCCTGGCCGCGGTGGTCTTGCTGGCCTTTGTCACCTGCTTCGCCCCC 971

Qy         336 SerAsnIleIleLeuIleIleHis---HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
               :::|||   :::|||:::   |||      :::      :::|||      :::
Db         972 AACAACTTCGTGCTCCTGGCGCACATCGTGAGCCGCCTGTTCTACGGCAAGAGC------ 1025

Qy         355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
               |||   :::|||   :::   |||||||||      ||||||:::||||||||||||
Db        1026 TACTACCACGTGTACAAGCTCACGCTGTGTCTCAGCTGCCTCAACAACTGTCTGGACCCG 1085

Qy         375 PheLeuTyrPheLeuMetSerLys 382
               |||:::|||:::      |||:::
Db        1086 TTTGTTTATTACTTTGCGTCCCGG 1109

RESULT 11
US-10-225-567A-529
; Sequence 529, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 529
;   LENGTH: 2137
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-529

Alignment Scores:
Pred. No.:                1.86e-33        Length:          2137
Score:                    465.00          Matches:         104
Percent Similarity:       51.62%          Conservative:    55
Best Local Similarity:    33.77%          Mismatches:      131
Query Match:              21.56%          Indels:          18
DB:                       9               Gaps:            8

US-09-208-629F-6 (1-408) x US-10-225-567A-529 (1-2137)

Qy          90 ProGluGluSerAlaSerHisLeuHisVal-----------------LysAsnAlaThr 103
               |||   |||      |||   :::   |||                 |||||||||
Db         195 CCGGGCGAGGACCCCTCCAGGATGCAGGTCCCGAACAGCACCGGCCCGGACAACGCGACG 254

Qy         104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
               :::      |||   :::                 :::|||   :::|||   ||||||
Db         255 CTGCAGATGCTGCGGAACCCGGCGATCGCGGTGGCCCTGCCCGTGGTGTACTCGCTGGTG 314

Qy         124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg------ 141
               |||   :::|||   |||      :::||||||:::|||      |||
Db         315 GCGGCGGTCAGCATCCCGGGCAACCTCTTCTCTCTGTGGGTGCTGTGCCGGCGCATGGGG 374
```



RESULT 12
US-09-739-151-1
; Sequence 1, Application US/09739151
; Patent No. US20010029032A1
; GENERAL INFORMATION:
; APPLICANT: Yuan Zhu
; APPLICANT: Xiaotong Li
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: PAUL, A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: POLYPEPTIDE
; FILE REFERENCE: GP-70567-C1
; CURRENT APPLICATION NUMBER: US/09/739,151
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/413,534
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/103,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-739-151-1

Alignment Scores:

Pred. No.:	2.06e-33	Length:	1080
Score:	460.00	Matches:	100
Percent Similarity:	52.92%	Conservative:	54
Best Local Similarity:	34.36%	Mismatches:	125
Query Match:	21.33%	Indels:	12
DB:	10	Gaps:	7

US-09-208-629F-6 (1-408) x US-09-739-151-1 (1-1080)

```
RESULT 13
US-10-251-385-181
; Sequence 181, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
;  TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
;  TITLE OF INVENTION:  Protein-Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/10/251,385
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US/09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 181
;   LENGTH: 1020
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-251-385-181

Alignment Scores:
Pred. No.:              3.88e-29      Length:          1020
Score:                  413.50        Matches:         108
Percent Similarity:     45.11%        Conservative:    49
Best Local Similarity:  31.03%        Mismatches:      130
Query Match:            19.17%        Indels:          61
DB:                     9             Gaps:            13

US-09-208-629F-6 (1-408) x US-10-251-385-181 (1-1020)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||||||||                |||       :::       |||
Db         19 GCTCCCCCAGGTCTGATCACCAACTTCTCCCTGGCCACGGCAGAG---------------- 63

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                                  :::|||   :::|||
Db         64 -----------------CAATGTGGCCAGGAG-------------------------- 78

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                                     :::   |||          |||      :::   ||||||
Db         79 --------------------ACGCCACTGGAGAACATGCTGTTCGCCTCCTTCTACCTT 117

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              |||   |||::::::  :::       |||   :::   ||||||   ||||||   |||
Db        118 CTGGATTTTATCCTGGCTTTAGTTGGCAATACCCTGGCTCTGTGG---CTTTTCATCCGA 174

Qy        142 ThrArgSerIleCysThr------ThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
                       |||          ||||||      :::|||||:::||||||
Db        175 GACCACAAGTCCGGGACCCCGGCCAACGTGTTCCTGATGCATCTGGCCGTGGCCGACTTG 234

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
                    :::   ||||||   ::::::   ||||||   :::||||||:::|||
Db        235 TCGTGCGTGCTGGTCCTGCCCACCCGCCTGGTCTACCACTTCTCTGGGAACCACTGGCCA 294

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
              |||||||||:::   ||||||   |||       :::||||||   |||||||||   |||
Db        295 TTTGGGGAAATCGCATGCCGTCTCACCGGCTTCCTCTTCTACCTCAACATGTACGCCAGC 354

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
              |||       |||   |||||||||   :::|||:::||||||||||||||||||
Db        355 ATCTACTTCCTCACCTGCATCAGCGCCGACCGTTTCCTGGCCATTGTGCACCCGGTCAAG 414

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
                       |||   :::       ||||||   :::   |||       :::|||
Db        415 TCCCTCAAGCTCCGCAGGCCCCTCTACGCACACCTGGCCTGTGCCTTCCTGTGGGTGGTG 474

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
              |||   :::   |||   |||       :::       |||

Db        475 GTGGCTGTGGCCATGGCCCCGCTGCTGGTGAGCCCACAG---------------------- 513

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
                 :::   |||   |||   |||          |||                   ||||||
Db        514 ACCGTGCAGACCAACCACACGGTG------GTCTGC--------------CTGCAGCTG 552

Qy        280 Tyr--------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              |||                        :::|||||||||         |||
Db        553 TACCGGGAGAAGGCCTCCCACCATGCCCTGGTGTCCCTGGCA---GTGGCCTTCACCTTC 609

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsn--------- 309
              ||||||:::       :::   ||||||      |||||||||:::|||
Db        610 CCGTTCATCACCACGGTCACCTGCTACCTGCTGATCATCCGCAGCCTGCGGCAGGGCCTG 669

Qy        310 AlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeu---LeuIleLeuValIle 328
                    :::   |||                |||       :::   ::::::|||   |||
Db        670 CGTGTGGAGAAGCGCCTCAAGACCAAGGCAAAACGCATGATCGCCATAGTGCTGGCCATC 729

Qy        329 PheThrIleCysPheAlaProSerAsnIle---------IleLeuIleIleHisHisAla 345
              |||   :::||||||   |||   ::::::          :::   ::::::|||:::
Db        730 TTCCTGGTCTGCTTCGTGCCCTACCACGTCAACCGCTCCGTCTACGTGCTGCACTACCGC 789

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGly------LeuTyrPheIleTyrLeuIleAlaLeu 363
              ::::::       :::                |||                   |||
Db        790 AGCCATGGGGCCTCCTGCGCCACCCAGCGCATCCTGGCCCTGGCAAACCGCATCACCTCC 849

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   |||||||||       |||||||||   :::||||||   :::::::::
Db        850 TGCCTCACCAGCCTCAACGGGGCACTCGACCCCATCATGTATTTCTTCGTGGCTGAGAAG 909

Qy        384 ArgAsnHisSerThrAlaTyrLeu 391
                    |||:::          |||
Db        910 TTCCGCCACGCCCTGTGCAACTTG 933

RESULT 14
US-10-251-385-31
; Sequence 31, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
;  TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
;  TITLE OF INVENTION:  Protein-Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/10/251,385
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US/09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
;   LENGTH: 1020
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-251-385-31

Alignment Scores:
Pred. No.:              7.36e-29      Length:          1020
Score:                  410.50        Matches:         107
Percent Similarity:     44.54%        Conservative:    48
Best Local Similarity:  30.75%        Mismatches:      132
Query Match:            19.03%        Indels:          61
DB:                     9             Gaps:            12

US-09-208-629F-6 (1-408) x US-10-251-385-31 (1-1020)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||||||||                |||       :::       |||
Db         19 GCTCCCCCAGGTCTGATCACCAACTTCTCCCTGGCCACGGCAGAG---------------- 63

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
```

```
Db         64 ------------------CAATGTGGCCAGGAG-------------------------- 78

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db         79 --------------------ACGCCACTGGAGAACATGCTGTTCGCCTCCTTCTACCTT 117

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db        118 CTGGATTTTATCCTGGCTTTAGTTGGCAATACCCTGGCTCTGTGG---CTTTTCATCCGA 174

Qy        142 ThrArgSerIleCysThr------ThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db        175 GACCACAAGTCCGGGACCCCGGCCAACGTGTTCCTGATGCATCTGGCCGTGGCCGACTTG 234

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db        235 TCGTGCGTGCTGGTCCTGCCCACCCGCCTGGTCTACCACTTCTCTGGGAACCACTGGCCA 294

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db        295 TTTGGGGAAATCGCATGCCGTCTCACCGGCTTCCTCTTCTACCTCAACATGTACGCCAGC 354

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db        355 ATCTACTTCCTCACCTGCATCAGCGCCGACCGTTTCCTGGCCATTGTGCACCCGGTCAAG 414

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db        415 TCCCTCAAGCTCCGCAGGCCCCTCTACGCACACCTGGCCTGTGCCTTCCTGTGGGTGGTG 474

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db        475 GTGGCTGTGGCCATGGCCCCGCTGCTGGTGAGCCCACAG---------------------- 513

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
Db        514 ACCGTGCAGACCAACCACACGGTG------GTCTGC---------------CTGCAGCTG 552

Qy        280 Tyr--------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db        553 TACCGGGAGAAGGCCTCCCACCATGCCCTGGTGTCCCTGGCA---GTGGCCTTCACCTTC 609

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr-----------Leu 308
Db        610 CCGTTCATCACCACGGTCACCTGCTACCTGCTGATCATCCGCAGCCTGCGGCAGGGCCTG 669

Qy        309 AsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db        670 CGTGTGGAGAAGCGCCTCAAGACCAAGGCAGTGCGCATGATCGCCATAGTGCTGGCCATC 729

Qy        329 PheThrIleCysPheAlaProSerAsnIle---------IleLeuIleIleHisHisAla 345
Db        730 TTCCTGGTCTGCTTCGTGCCCTACCACGTCAACCGCTCCGTCTACGTGCTGCACTACCGC 789

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGly------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db        790 AGCCATGGGGCCTCCTGCGCCACCCAGCGCATCCTGGCCCTGGCAAACCGCATCACCTCC 849

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db        850 TGCCTCACCAGCCTCAACGGGGCACTCGACCCCATCATGTATTTCTTCGTGGCTGAGAAG 909

Qy        384 ArgAsnHisSerThrAlaTyrLeu 391
Db        910 TTCCGCCACGCCCTGTGCAACTTG 933

RESULT 15
US-09-788-133-1
; Sequence 1, Application US/09788133
; Patent No. US20020052001A1
; GENERAL INFORMATION:
;  APPLICANT: GLAXO GROUP LTD
;  TITLE OF INVENTION: ASSAY
;  FILE REFERENCE: P79011
;  CURRENT APPLICATION NUMBER: US/09/788,133
;  CURRENT FILING DATE:  2001-02-20
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1020
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1020)
US-09-788-133-1

Alignment Scores:
Pred. No.:              7.36e-29        Length:         1020
Score:                  410.50          Matches:        107
Percent Similarity:     44.54%          Conservative:   48
Best Local Similarity:  30.75%          Mismatches:     132
Query Match:            19.03%          Indels:         61
DB:                     10              Gaps:           12

US-09-208-629F-6 (1-408) x US-09-788-133-1 (1-1020)

Qy         62 AlaProProAsnSerPheGluGluProPheProSerAlaLeuGluGlyTrpThrGlyAla 81
Db         19 GCTCCCCCAGGTCTGATCACCAACTTCTCCCTGGCCACGGCAGAG--------------- 63

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db         64 ------------------CAATGTGGCCAGGAG-------------------------- 78

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db         79 --------------------ACGCCACTGGAGAACATGCTGTTCGCCTCCTTCTACCTT 117

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db        118 CTGGATTTTATCCTGGCTTTAGTTGGCAATACCCTGGCTCTGTGG---CTTTTCATCCGA 174

Qy        142 ThrArgSerIleCysThr------ThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db        175 GACCACAAGTCCGGGACCCCGGCCAACGTGTTCCTGATGCATCTGGCCGTGGCCGACTTG 234

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db        235 TCGTGCGTGCTGGTCCTGCCCACCCGCCTGGTCTACCACTTCTCTGGGAACCACTGGCCA 294

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db        295 TTTGGGGAAATCGCATGCCGTCTCACCGGCTTCCTCTTCTACCTCAACATGTACGCCAGC 354

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db        355 ATCTACTTCCTCACCTGCATCAGCGCCGACCGTTTCCTGGCCATTGTGCACCCGGTCAAG 414

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db        415 TCCCTCAAGCTCCGCAGGCCCCTCTACGCACACCTGGCCTGTGCCTTCCTGTGGGTGGTG 474

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db        475 GTGGCTGTGGCCATGGCCCCGCTGCTGGTGAGCCCACAG---------------------- 513

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
Db        514 ACCGTGCAGACCAACCACACGGTG------GTCTGC---------------CTGCAGCTG 552

Qy        280 Tyr--------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db        553 TACCGGGAGAAGGCCTCCCACCATGCCCTGGTGTCCCTGGCA---GTGGCCTTCACCTTC 609
```

```
Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr----------Leu 308
Db        610 CCGTTCATCACCACGGTCACCTGCTACCTGCTGATCATCCGCAGCCTGCGGCAGGGCCTG 669

Qy        309 AsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db        670 CGTGTGGAGAAGCGCCTCAAGACCAAGGCAGTGCGCATGATCGCCATAGTGCTGGCCATC 729

Qy        329 PheThrIleCysPheAlaProSerAsnIle---------IleLeuIleIleHisHisAla 345
Db        730 TTCCTGGTCTGCTTCGTGCCCTACCACGTCAACCGCTCCGTCTACGTGCTGCACTACCGC 789

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGly------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db        790 AGCCATGGGGCCTCCTGCGCCACCCAGCGCATCCTGGCCCTGGCAAACCGCATCACCTCC 849

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db        850 TGCCTCACCAGCCTCAACGGGGCACTCGACCCCATCATGTATTTCTTCGTGGCTGAGAAG 909

Qy        384 ArgAsnHisSerThrAlaTyrLeu 391
Db        910 TTCCGCCACGCCCTGTGCAACTTG 933
```

RESULT 16
US-10-222-024-1
; Sequence 1, Application US/10222024
; Publication No. US20030104487A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Ltd. (GB)
; APPLICANT: Pfizer Inc. (EP except GB, US, JP)
; TITLE OF INVENTION: Neuropeptide receptor and uses thereof
; FILE REFERENCE: PCS22032
; CURRENT APPLICATION NUMBER: US/10/222,024
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: GB 0119920.7
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-222-024-1

Alignment Scores:

Pred. No.:	8.04e-29	Length:	1086
Score:	410.50	Matches:	99
Percent Similarity:	47.87%	Conservative:	58
Best Local Similarity:	30.18%	Mismatches:	114
Query Match:	19.03%	Indels:	57
DB:	9	Gaps:	10

US-09-208-629F-6 (1-408) x US-10-222-024-1 (1-1086)

```
Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
Db         70 TATGCACATCACAGCACGGCCAGGATAGTAATGCCTCTGCATTACAGCCTCGTCTTCATC 129

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db        130 ATTGGGCTCGTGGGAAACTTACTAGCCTTGGTCGTCATTGTTCAAAACAGGAAAAAAATC 189

Qy        146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
Db        190 AACTCTACCACCCTCTATTCAACAAATTTGGTGATTTCTGATATACTTTTTACCACGGCT 249

Qy        165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
Db        250 TTGCCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGAGAATCGGAGATGCCTTG 309

Qy        185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaLeu 204
Db        310 TGTAGGATAACTGCGCTAGTGTTTTACATCAACACATATGCAGGTGTGAACTTTATGACC 369

Qy        205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224
Db        370 TGCCTGAGTATTGACCGCTTCATTGCTGTGGTGCACCCTCTACGCTACAACAAGATAAAA 429

Qy        225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
Db        430 AGGATTGAACATGCAAAAGGCGTGTGCATATTTGTCTGGATTCTAGTATTTGCTCAGACA 489

Qy        245 LeuProPhePheIle---------LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db        490 CTCCCACTCCTCATCAACCCTATGTCAAAGCAGGAG----------GCTGAAAGG 534

Qy        262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer---ProPheGlnLeuTyr 280
Db        535 ATTACATGCATGGAGTATCCAAACTTTGAAGAAACTAAATCTCTTCCCTGGATTCTG--- 591

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db        592 ------CTTGGGGCATGTTTCATAGGATATGTACTTCCACTTATAATCATTCTCATCTGC 645

Qy        301 TyrAlaAlaIleIle---------------------------------------------- 305
Db        646 TATTCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAAACCCACTCACTGAGAAA 705

Qy        306 ---------------ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
Db        706 TCTGGTGTAAACAAAAAGGCTCTCAACACAATT------------------------- 738

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
Db        739 ------ATTCTTATTATTGTTGTGTTTGTTCTCTGTTTCACACCTTACCATGTTGCAATT 792

Qy        341 IleIleHis-----------------HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db        793 ATTCAACATATGATTAAGAAGCTTCGTTTCTCTAATTTCCTGGAATGTAGCCAAAGACAT 852

Qy        355 LeuTyrPheIleTyrLeu---IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
Db        853 TCGTTCCAGATTTCTCTGCACTTTACAGTATGCCTGATGAACTTCAATTGCTGCATGGAC 912

Qy        374 ProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys 393
Db        913 CCTTTTATCTACTTCTTTGCATGTAAA-----------------GGGTATAAGAGAAAG 954

Qy        394 ******AsnAspLeuArgGluGln 401
Db        955 GTTATGAGGATGCTGAAACGGCAA 978
```

RESULT 17
US-10-251-385-77
; Sequence 77, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-77

```
Alignment Scores:
Pred. No.:                8.04e-29        Length:          1086
Score:                    410.50          Matches:         99
Percent Similarity:       47.87%          Conservative:    58
Best Local Similarity:    30.18%          Mismatches:      114
Query Match:              19.03%          Indels:          57
DB:                       9               Gaps:            10

US-09-208-629F-6 (1-408) x US-10-251-385-77 (1-1086)

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
              |||         |||   :::      ::::::|||      |||   |||||||||:::
Db         70 TATGCACATCACAGCACGGCCAGGATAGTAATGCCTCTGCATTACAGCCTCGTCTTCATC 129

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
              :::|||:::      |||   :::   |||   ::::::             :::   |||
Db        130 ATTGGGCTCGTGGGAAACTTACTAGCCTTGGTCGTCATTGTTCAAAACAGGAAAAAAATC 189

Qy        146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
                 ||||||::::::  |||||||||   |||:::|||   ||||||
Db        190 AACTCTACCACCCTCTATTCAACAAATTTGGTGATTTCTGATATACTTTTTACCACGGCT 249

Qy        165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
              ||||||   :::|||||||||:::      |||   :::|||      |||:::   |||
Db        250 TTGCCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGAGAATCGGAGATGCCTTG 309

Qy        185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAla 204
              ||||||   |||   ::::::||||||   |||   |||       :::       :::
Db        310 TGTAGGATAACTGCGCTAGTGTTTTACATCAACACATATGCAGGTGTGAACTTTATGACC 369

Qy        205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224
              |||:::||||||:::|||::::::|||:::|||||||||      |||      :::
Db        370 TGCCTGAGTATTGACCGCTTCATTGCTGTGGTGCACCCTCTACGCTACAACAAGATAAAA 429

Qy        225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
              :::       :::|||          |||      ||||||      ||||||
Db        430 AGGATTGAACATGCAAAAGGCGTGTGCATATTTGTCTGGATTCTAGTATTTGCTCAGACA 489

Qy        245 LeuProPhePheIle---------LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
              ||||||          |||                |||||||||                :::
Db        490 CTCCCACTCCTCATCAACCCTATGTCAAAGCAGGAG--------------GCTGAAAGG 534

Qy        262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer---ProPheGlnLeuTyr 280
              ||||||   :::      |||      |||:::   |||   |||:::   |||
Db        535 ATTACATGCATGGAGTATCCAAACTTTGAAGAAACTAAATCTCTTCCCTGGATTCTG--- 591

Qy        281 TyrPheIleSerLeuAlaPheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
                  :::          |||   |||:::::::::|||   ::::::|||:::   |||
Db        592 ------CTTGGGGCATGTTTCATAGGATATGTACTTCCACTTATAATCATTCTCATCTGC 645

Qy        301 TyrAlaAlaIleIle---------------------------------------------- 305
              |||:::   |||
Db        646 TATTCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAAACCCACTCACTGAGAAA 705

Qy        306 ---------------ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
                            :::   ||||||
Db        706 TCTGGTGTAAACAAAAAGGCTCTCAACACAATT-------------------------- 738

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
                  :::||||||:::|||:::|||   :::|||||   |||   :::::   :::
Db        739 ------ATTCTTATTATTGTTGTGTTTGTTCTCTGTTTCACACCTTACCATGTTGCAATT 792

Qy        341 IleIleHis-----------------HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
              |||   |||                     :::|||:::      :::
Db        793 ATTCAACATATGATTAAGAAGCTTCGTTTCTCTAATTTCCTGGAATGTAGCCAAAGACAT 852

Qy        355 LeuTyrPheIleTyrLeu---IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
                 :::   |||   |||       :::||||||   :::   |||   |||:::|||
Db        853 TCGTTCCAGATTTCTCTGCACTTTACAGTATGCCTGATGAACTTCAATTGCTGCATGGAC 912

Qy        374 ProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys 393
              ||||||:::||||||         |||                         |||   |||
```

```
Db        913 CCTTTTATCTACTTCTTTGCATGTAAA----------------GGGTATAAGAGAAAG 954

Qy        394 ******AsnAspLeuArgGluGln 401
                      |||:::   |||
Db        955 GTTATGAGGATGCTGAAACGGCAA 978

RESULT 18
US-10-251-385-205
; Sequence 205, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
;  TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
;  TITLE OF INVENTION:  Protein-Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/10/251,385
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US/09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 205
;   LENGTH: 1086
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-251-385-205

Alignment Scores:
Pred. No.:                8.04e-29        Length:          1086
Score:                    410.50          Matches:         102
Percent Similarity:       50.47%          Conservative:    58
Best Local Similarity:    32.18%          Mismatches:      122
Query Match:              19.03%          Indels:          35
DB:                       9               Gaps:            11

US-09-208-629F-6 (1-408) x US-10-251-385-205 (1-1086)

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
              |||         |||   :::      ::::::|||      |||   |||||||||:::
Db         70 TATGCACATCACAGCACGGCCAGGATAGTAATGCCTCTGCATTACAGCCTCGTCTTCATC 129

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
              :::|||:::      |||   :::   |||   ::::::             :::   |||
Db        130 ATTGGGCTCGTGGGAAACTTACTAGCCTTGGTCGTCATTGTTCAAAACAGGAAAAAAATC 189

Qy        146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
                 ||||||::::::  |||||||||   |||:::|||   ||||||
Db        190 AACTCTACCACCCTCTATTCAACAAATTTGGTGATTTCTGATATACTTTTTACCACGGCT 249

Qy        165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
              ||||||   :::|||||||||:::      |||   :::|||      |||:::   |||
Db        250 TTGCCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGAGAATCGGAGATGCCTTG 309

Qy        185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAla 204
              ||||||   |||   ::::::||||||   |||   |||       :::       :::
Db        310 TGTAGGATAACTGCGCTAGTGTTTTACATCAACACATATGCAGGTGTGAACTTTATGACC 369

Qy        205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224
              |||:::||||||:::|||::::::|||:::|||||||||      |||      :::
Db        370 TGCCTGAGTATTGACCGCTTCATTGCTGTGGTGCACCCTCTACGCTACAACAAGATAAAA 429

Qy        225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
              :::       :::|||          |||      ||||||      ||||||
Db        430 AGGATTGAACATGCAAAAGGCGTGTGCATATTTGTCTGGATTCTAGTATTTGCTCAGACA 489

Qy        245 LeuProPhePheIle---------LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
              ||||||          |||                |||||||||                :::
Db        490 CTCCCACTCCTCATCAACCCTATGTCAAAGCAGGAG--------------GCTGAAAGG 534

Qy        262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer---ProPheGlnLeuTyr 280
```

```
Db        535 ATTACATGCATGGAGTATCCAAACTTTGAAGAAACTAAATCTCTTCCCTGGATTCTG--- 591

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300

Db        592 ------CTTGGGGCATGTTTCATAGGATATGTACTTCCACTTATAATCATTCTCATCTGC 645

Qy        301 TyrAlaAlaIle------------IleArgThrLeu------AsnAlaTyrAspHisArg 314

Db        646 TATTCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAAACCCACTCACTGAGAAA 705

Qy        315 TrpLeuTrpTyrValLysAlaSer---------LeuLeuIleLeuValIlePheThrIle 331

Db        706 TCTGGTGTAAACAAAAAGGCTAAAAACACAATTATTCTTATTATTGTTGTGTTTGTTCTC 765

Qy        332 CysPheAlaProSerAsnIleIleLeuIleIleHis-----------------HisAla 345

Db        766 TGTTTCACACCTTACCATGTTGCAATTATTCAACATATGATTAAGAAGCTTCGTTTCTCT 825

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu---IleAlaLeuCys 364

Db        826 AATTTCCTGGAATGTAGCCAAAGACATTCGTTCCAGATTTCTCTGCACTTTACAGTATGC 885

Qy        365 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThrArg 384

Db        886 CTGATGAACTTCAATTGCTGCATGGACCCTTTTATCTACTTCTTTGCATGTAAA------ 939

Qy        385 AsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGluGln 401

Db        940 -----------GGGTATAAGAGAAAGGTTATGAGGATGCTGAAACGGCAA 978

RESULT 19
US-10-225-567A-111
; Sequence 111, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 111
;   LENGTH: 1638
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-111

Alignment Scores:
Pred. No.:              1.43e-28        Length:         1638
Score:                  410.50          Matches:        99
Percent Similarity:     47.87%          Conservative:   58
Best Local Similarity:  30.18%          Mismatches:     114
Query Match:            19.03%          Indels:         57
DB:                     9               Gaps:           10

US-09-208-629F-6 (1-408) x US-10-225-567A-111 (1-1638)

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125

Db        103 TATGCACATCACAGCACGGCCAGGATAGTAATGCCTCTGCATTACAGCCTCGTCTTCATC 162

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145

Db        163 ATTGGGCTCGTGGGAAACTTACTAGCCTTGGTCGTCATTGTTCAAAACAGGAAAAAAATC 222

Qy        146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164

Db        223 AACTCTACCACCCTCTATTCAACAAATTTGGTGATTTCTGATATACTTTTTACCACCGCT 282

Qy        165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184

Db        283 TTGCCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGAGAATCGGAGATGCCTTG 342

Qy        185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAla 204

Db        343 TGTAGGATAACTGCGCTAGTGTTTTACATCAACACATATGCAGGTGTGAACTTTATGACC 402

Qy        205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224

Db        403 TGCCTGAGTATTGACCGCTTCATTGCTGTGGTGCACCCTCTACGCTACAACAAGATAAAA 462

Qy        225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244

Db        463 AGGATTGAACATGCAAAAGGCGTGTGCATATTTGTCTGGATTCTAGTATTTGCTCAGACA 522

Qy        245 LeuProPhePheIle---------LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261

Db        523 CTCCCACTCCTCATCAACCCTATGTCAAAGCAGGAG--------------GCTGAAAGG 567

Qy        262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer---ProPheGlnLeuTyr 280

Db        568 ATTACATGCATGGAGTATCCAAACTTTGAAGAAACTAAATCTCTTCCCTGGATTCTG--- 624

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300

Db        625 ------CTTGGGGCATGTTTCATAGGATATGTACTTCCACTTATAATCATTCTCATCTGC 678

Qy        301 TyrAlaAlaIleIle-------------------------------------------- 305

Db        679 TATTCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAAACCCACTCACTGAGAAA 738

Qy        306 --------------ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320

Db        739 TCTGGTGTAAACAAAAAGGCTCTCAACACAATT-------------------------- 771

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340

Db        772 ------ATTCTTATTATTGTTGTGTTTGTTCTCTGTTTCACACCTTACCATGTTGCAATT 825

Qy        341 IleIleHis-----------------HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354

Db        826 ATTCAACATATGATTAAGAAGCTTCGTTTCTCTAATTTCCTGGAATGTAGCCAAAGACAT 885

Qy        355 LeuTyrPheIleTyrLeu---IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373

Db        886 TCGTTCCAGATTTCTCTGCACTTTACAGTATGCCTGATGAACTTCAATTGCTGCATGGAC 945

Qy        374 ProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys 393

Db        946 CCTTTTATCTACTTCTTTGCATGTAAA-----------------GGGTATAAGAGAAAG 987

Qy        394 ******AsnAspLeuArgGluGln 401

Db        988 GTTATGAGGATGCTGAAACGGCAA 1011

RESULT 20
US-10-225-567A-303
; Sequence 303, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;. TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
```

```
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 303
;   LENGTH: 2070
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-303

Alignment Scores:
Pred. No.:                1.99e-28        Length:          2070
Score:                    410.50          Matches:         107
Percent Similarity:       44.54%          Conservative:    48
Best Local Similarity:    30.75%          Mismatches:      132
Query Match:              19.03%          Indels:          61
DB:                       9               Gaps:            12

US-09-208-629F-6 (1-408) x US-10-225-567A-303 (1-2070)

Qy         62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
              |||||||||               |||      :::      |||
Db        177 GCTCCCCCAGGTCTGATCACCAACTTCTCCCTGGCCACGGCAGAG---------------- 221

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
                          :::|||   :::|||
Db        222 -----------------CAATGTGGCCAGGAG-------------------------- 236

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                         :::   |||          |||      :::   ||||||
Db        237 --------------------ACGCCACTGGAGAACATGCTGTTCGCCTCCTTCTACCTT 275

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              |||   |||::::::  :::      |||   :::   ||||||   ||||||   |||
Db        276 CTGGATTTTATCCTGGCTTTAGTTGGCAATACCCTGGCTCTGTGG---CTTTTCATCCGA 332

Qy        142 ThrArgSerIleCysThr------ThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
                       |||         ||||||      :::|||||:::||||||
Db        333 GACCACAAGTCCGGGACCCCGGCCAACGTGTTCCTGATGCATCTGGCCGTGGCCGACTTG 392

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
                  :::   ||||||   ::::::   ||||||   :::|||||:::|||
Db        393 TCGTGCGTGCTGGTCCTGCCCACCCGCCTGGTCTACCACTTCTCTGGGAACCACTGGCCA 452

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
              |||||||||:::   ||||||   |||      :::||||||   |||||||||   |||
Db        453 TTTGGGGAAATCGCATGCCGTCTCACCGGCTTCCTCTTCTACCTCAACATGTACGCCAGC 512

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
              |||      |||   |||||||||   :::|||:::||||||||||||||||||||
Db        513 ATCTACTTCCTCACCTGCATCAGCGCCGACCGTTTCCTGGCCATTGTGCACCCGGTCAAG 572

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
                    |||   :::      ||||||   :::   |||      :::|||
Db        573 TCCCTCAAGCTCCGCAGGCCCCTCTACGCACACCTGGCCTGTGCCTTCCTGTGGGTGGTG 632

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
              |||   :::   |||   |||      :::      |||
Db        633 GTGGCTGTGGCCATGGCCCCGCTGCTGGTGAGCCCACAG---------------------- 671

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
                 :::   |||   |||   |||         |||                  ||||||
Db        672 ACCGTGCAGACCAACCACACGGTG------GTCTGC---------------CTGCAGCTG 710

Qy        280 Tyr--------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              |||                          :::|||||||||         |||
Db        711 TACCGGGAGAAGGCCTCCCACCATGCCCTGGTGTCCCTGGCA---GTGGCCTTCACCTTC 767

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr-----------Leu 308
              ||||||:::      :::  ||||||      |||||||||:::             |||
Db        768 CCGTTCATCACCACGGTCACCTGCTACCTGCTGATCATCCGCAGCCTGCGGCAGGGCCTG 827

Qy        309 AsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
                                       |||:::        ::::::|||   |||
```

```
Db        828 CGTGTGGAGAAGCGCCTCAAGACCAAGGCAGTGCGCATGATCGCCATAGTGCTGGCCATC 887

Qy        329 PheThrIleCysPheAlaProSerAsnIle---------IleLeuIleIleHisHisAla 345
              |||   :::||||||   |||   ::::::         :::   :::::::|||:::
Db        888 TTCCTGGTCTGCTTCGTGCCCTACCACGTCAACCGCTCCGTCTACGTGCTGCACTACCGC 947

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGly------LeuTyrPheIleTyrLeuIleAlaLeu 363
              ::::::      :::                |||              |||
Db        948 AGCCATGGGGCCTCCTGCGCCACCCAGCGCATCCTGGCCCTGGCAAACCGCATCACCTCC 1007

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   |||||||||       |||||||||   :::||||||   :::::::::
Db       1008 TGCCTCACCAGCCTCAACGGGGCACTCGACCCCATCATGTATTTCTTCGTGGCTGAGAAG 1067

Qy        384 ArgAsnHisSerThrAlaTyrLeu 391
                   |||:::        |||
Db       1068 TTCCGCCACGCCCTGTGCAACTTG 1091

RESULT 21
US-10-225-567A-218
; Sequence 218, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
;   LENGTH: 2424
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-218

Alignment Scores:
Pred. No.:                1.7e-27         Length:          2424
Score:                    401.50          Matches:         123
Percent Similarity:       46.62%          Conservative:    70
Best Local Similarity:    29.71%          Mismatches:      161
Query Match:              18.61%          Indels:          60
DB:                       9               Gaps:            15

US-09-208-629F-6 (1-408) x US-10-225-567A-218 (1-2424)

Qy         20 MetLysAlaLeuIlePheAlaAlaAlaGlyLeu-LeuLeuLeuLeuProThrPheCysGl 39
              |||   |||          |||||||||   |||       |||||||||
Db         56 ATGCTTGCTGCGCCCCTGGCCGCCGCTGCCCTCTCGCCGCCTCCTACCCCTCGGAGCCGC 115

Qy         39 nSerGlyMetGluAsnAsp---ThrAsnAsnLeuAlaLysProThrLeuProIleLysTh 58
               :::   :::|||   :::             :::      |||   :::|||      ||
Db        116 CGCCTAAGTCGAGGAGGAGAGAATGACCGAGGTGCTGTGGCCGGCTGTCCCCAACGGGAC 175

Qy         58 r---------PheArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLe 75
              |         |||          |||   :::|||
Db        176 GGACGCTGCCTTCCTGGCCGGTCCGGGTTCGTCC------------------------- 209

Qy         75 uGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSe 95
                    |||       :::|||:::                              |||||
Db        210 -------TGGGGGAACAGCACGGTC----------------------------GCCTC 232

Qy         95 rHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr---SerSerLeuSerThrLy 114
              |         |||   ::::::                |||||   :::
Db        233 CACTGCCGCCGTCTCCTCGTCGTTCAAATGCGCCTTGACCAAGACGGGCTTCCAGTTTTA 292

Qy        114 sLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValTh 134
                 :::||||||:::|||:::|||||||||:::::|||         |||:::|||
```

```
Db        293 CTACCTGCCGGCTGTCTACATCTTGGTATTCATCATCGGCTTCCTGGGCAACAGCGTGGC 352

Qy        134 rLeuTrpMetLeuPhePheArgThrArgSerIleCys---ThrThrValPheTyrThrAs 153
              :::||||||      |||      :::              :::|||:::       ||
Db        353 CATCTGGATGTTCGTCTTCCACATGAAGCCCTGGAGCGGCATCTCCGTGTACATGTTCAA 412

Qy        153 nLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLe 173
              |||||||:::|||||||||||||:::   :::|||||||||      |||   |||:::
Db        413 TTTGGCTCTGGCCGACTTCTTGTACGTGCTGACTCTGCCAGCCCTGATCTTCTACTACTT 472

Qy        173 uAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTy 193
              |||      :::|||:::||||||:::   :::|||:::            ||||||::
Db        473 CAATAAAACAGACTGGATCTTCGGGGATGCCATGTGTAAACTGCAGAGGTTCATCTTTCA 532

Qy        193 rGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAl 213
              :   |||:::|||   |||||||||   |||   |||||||||   :::||||||
Db        533 TGTGAACCTCTATGGCAGCATCTTGTTTCTGACATGCATCAGTGCCCACCGGTACAGCGG 592

Qy        213 aIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCy 233
              :::|||:::|||          |||      |||          |||:::
Db        593 TGTGGTGTACCCCCTCAAGTCCCTGGGCCGGCTCAAAAAGAAGAATGCGATCTGTATCAG 652

Qy        233 sGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGl 253
              |||||||||      |||   :::   :::   |||      ||||||
Db        653 CGTGCTGGTGTGGCTCATTGTGGTGGTGGCGATCTCCCCC------ATCCTCTTCTACTC 706

Qy        253 uTyrTyrLeuValGlnProAspIleThr---ThrCysHisAspValHisAsnThrCysGl 272
                  |||:::   :::   |||   ||||||:::|||
Db        707 AGGTACCGGGGTCCGCAAAAACAAAACCATCACCTGTTACGACACC-------------- 752

Qy        272 uSerSerSerProPheGlnLeuTyrTyrPheIle--------------SerLeuAlaPh 287
              :::|||      :::          |||||||||                :::::|||
Db        753 -ACCTCAGACGAGTACCTGCGAAGTTATTTCATCTACAGCATGTGCACGACCGTGGCCAT 811

Qy        287 ePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgTh 307
                  |||   :::|||   |||||||||:::   ||||||      |||:::|||
Db        812 G------TTCTGTGTCCCCTTGGTGCTGATTCTGGGCTGTTACGGATTAATTGTGAGAGC 865

Qy        307 rLeuAsnAlaTyrAsp---------------------HisArgTrpLeuTrpTyrValLy 320
              |||          |||                    :::   ::::::   |||
Db        866 TTTGATTTACAAAGATCTGGACAACTCTCCTCTGAGGAGAAAATCGATTTACCTGGTA-- 923

Qy        320 sAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLe 340
              :::::::::|||   :::|||   :::   :::   |||   :::::::::
Db        924 -------ATCATTGTACTGACTGTTTTTGCTGTGTCTTACATCCCTTTCCATGTGATGAA 976

Qy        340 uIleIleHis---HisAlaAsnTyrTyrTyrAsnAsn-----------------ThrAs 353
                  ::::::      |||          :::                        ||
Db        977 AACGATGAACTTGAGGGCCCGGCTTGATTTTCAGACCCCAGCAATGTGTGCTTTCAATGA 1036

Qy        353 pGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAs 373
              |   :::|||      |||   :::          |||   |||||||||||||:::||
Db       1037 CAGGGTTTATGCCACGTATCAGGTGACAAGAGGTCTAGCAAGTCTCAACAGTTGTGTGGA 1096

Qy        373 pProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLy 393
              ||||   ||||||||||||          |||                :::      ||::
Db       1097 CCCCATTCTCTATTTCTTGGCGGGAGATACTTTCAGAAGGAGACTCTCCCGAGCCACAAG 1156

Qy        393 s******AsnAspLeuArgGluGlnGlyGlnProSerGln 406
              :      :::          |||             |||:::
Db       1157 GAAAGCTTCTAGAAGAAGTGAGGCAAATTTGCAATCCAAG 1196

RESULT 22
US-10-024-494-7
; Sequence 7, Application US/10024494
; Publication No. US20030044898A1
;   GENERAL INFORMATION:
;        APPLICANT: LI, YI
;                   CAO, LIANG
;                   ROSEN, CRAIG A.
;        TITLE OF INVENTION: Human GPR4 G-Protein Coupled Receptor and
;                            Nucleotides Encoding Same (As Amended)
;        NUMBER OF SEQUENCES: 30
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;             STREET: 1100 NEW YORK AVE., NW, SUITE 600
;             CITY: WASHINGTON
;             STATE: DC
;             COUNTRY: USA
;             ZIP: 20005
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: FLOPPY DISK
;             COMPUTER: IBM PC COMPATIBLE
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/024,494
;             FILING DATE: 21-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/465,973
;             FILING DATE: 06-JUN-1995
;             APPLICATION NUMBER: PCT/US95/04079
;             FILING DATE: 30-MAR-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: STEFFE, ERIC K.
;             REGISTRATION NUMBER: 36,688
;             REFERENCE/DOCKET NUMBER: 1488.1140001/EKS/HCC
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 202-371-2600
;             TELEFAX: 202-371-2540
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1301 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: both
;             TOPOLOGY: both
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  161..1192
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-024-494-7

Alignment Scores:
Pred. No.:              1.02e-26      Length:          1301
Score:                  389.00        Matches:         105
Percent Similarity:     47.69%        Conservative:    60
Best Local Similarity:  30.35%        Mismatches:      130
Query Match:            18.03%        Indels:          51
DB:                     9             Gaps:            11

US-09-208-629F-6 (1-408) x US-10-024-494-7 (1-1301)

Qy         77 GlyTrpThrGlyAlaThrIleThrValLysIleLysCys-ProGluGluSerAlaSerHi 96
              |||:::   |||      :::      |||      |||    :::|||||||||||||
Db         76 GGTTATAGAGGTTATAATCTGAATCCCAAAGGAGACTGCAGCTGATGAAAGTGCTTCCAA 135

Qy         96 sLeuHisVal----------------------LysAsnAlaThrMetGlyTyrLeuTh 108
              |||   :::                        |||::::::      :::
Db        136 ACTGAAAATTGGACGTGCCTTTACGATGGTAAGCGTTAACAGCTCCCACTGCTTCTATAA 195

Qy        108 rSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVa 128
                 |||          |||      ::::::   :::|||||||||:::|||::
Db        196 TGACTCCTTTAAGTACACTTTGTATGGGTGCATGTTCAGCATGGTGTTTGTGCTTGGGTT 255

Qy        128 lProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys----- 146
              :   :::|||   |||   :::::::::                      ||||||
Db        256 AATATCCAATTGTGTTGCCATATACATTTTC----------------ATCTGCGTCCT 297

Qy        147 ---------------ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPh 161
                              ||||||   :::      |||||||||:::::|||   |||||
Db        298 CAAAGTCCGAAATGAAACTACAACTTACATGATTAACTTGGCAATGTCAGACTTGCTTTT 357

Qy        161 eCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGl 181
              |      ||||||||||||:::|||   ::::::            ||||||   |||||
```

```
Db        358 TGTTTTTACTTTACCCTTCAGGATT---TTTTACTTCACAACACGGAATTGGCCATTTGG 414

Qy        181 yGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLe 201

Db        415 AGATTTACTTTGTAAGATTTCTGTGATGCTGTTTTATACCAACATGTACGGAAGCATTCT 474

Qy        201 uLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrAr 221

Db        475 GTTCTTAACCTGTATTAGTGTAGATCGATTTCTGGCAATTGTCTACCCATTTAAGTCAAA 534

Qy        221 gGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPh 241

Db        535 GACTCTAAGAACCAAAAGAAATGCAAAGATTGTTTGCACTGGCGTGTGGTTAACTGTGAT 594

Qy        241 eLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIl 261

Db        595 CGGAGGAAGTGCACCC--------------------GCCGTTTTTGTTCAGTCTACCCA 633

Qy        261 eThrThrCysHisAspValHisAsnThrCysGluSerSerSerPro-------------- 276

Db        634 CTCTCAGGGTAACAATGCCTCAGAAGCCTGCTTTGAAAATTTTCCAGAAGCCACATGGAA 693

Qy        277 ----------PheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIlePr 293

Db        694 AACATATCTCTCAAGGATTGTAATTTTCATCGAAATAGTG------GGATTTTTTATTCC 747

Qy        293 oPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsn---------Al 310

Db        748 TCTAATTTTAAATGTAACTTGTTCTAGTATGGTGCTAAAAACTTTAACCAAACCTGTTAC 807

Qy        310 aTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle---------LeuVa 327

Db        808 ATTAAGTAGAAGCAAAATAAACAAAACTAAGGTTTTAAAAATGATTTTTGTACATTTGAT 867

Qy        327 lIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHis---------Hi 344

Db        868 CATATTCTGTTTCTGTTTTGTTCCTTACAATATCAATCTTATTTTATATTCTCTTGTGAG 927

Qy        344 sAlaAsnTyrTyrTyrAsnAsnThr------AspGlyLeuTyrPheIleTyrLeuIleAl 362

Db        928 AACACAAACATTTGTTAATTGCTCAGTAGTGGCAGCAGTAAGGACAATGTACCCAATCAC 987

Qy        362 aLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLy 382

Db        988 TCTCTGTATTGCTGTTTCCAACTGTTGTTTTGACCCTATAGTTTACTACTTTACATCGGA 1047

Qy        382 sThrArgAsnHisSer 387

Db       1048 CACAATTCAGAATTCA 1063

RESULT 23
US-10-121-101B-9/c
; Sequence 9, Application US/10121101B
; Publication No. US20030082653A1
; GENERAL INFORMATION:
;  APPLICANT: Lal, Preeti G.
;  APPLICANT:  Jackson, Jennifer L.
;  APPLICANT:  Krasnow, Randi
;  APPLICANT:  Rickert, Paula
;  TITLE OF INVENTION: GPCR Diferentially Expressed in Squamous Cell Carcinoma
;  FILE REFERENCE: PV-0004 CIP
;  CURRENT APPLICATION NUMBER: US/10/121,101B
;  CURRENT FILING DATE:  2002-04-11
;  PRIOR APPLICATION NUMBER: 09/470,252
;  PRIOR FILING DATE: 1999-12-22
;  PRIOR APPLICATION NUMBER: 08/988,876
;  PRIOR FILING DATE: 1997-12-11
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PERL Program
; SEQ ID NO 9
;   LENGTH: 1163
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030082653A1 093983_Mm.1
US-10-121-101B-9

Alignment Scores:
Pred. No.:              1.33e-26        Length:         1163
Score:                  387.00          Matches:        100
Percent Similarity:     47.89%          Conservative:   70
Best Local Similarity:  28.17%          Mismatches:     131
Query Match:            17.94%          Indels:         54
DB:                     9               Gaps:           11

US-09-208-629F-6 (1-408) x US-10-121-101B-9 (1-1163)

Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGlu 68

Db       1163 ATGGCTGTCCCTAATGTGAATGTTTCTACTTTTGCAGGTAAT------------------ 1122

Qy         69 GluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLys 88

Db       1121 --------------------GAGCTGTACAGCCAAGCAAGTCACACTGCA--------- 1092

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108

Db       1091 ---------AACAGCACAAGTGAGGGACATGGGAAGAACTCTACC---------CTGCAC 1050

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128

Db       1049 AACAAATTTGACACCATCATCCTGCCAGTGCTTTACCTAGTTATCTTTGTGGCAAGCATC 990

Qy        129 ProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThr 148

Db        989 CTGCTGAACGGTCTGGCCGTGTGG---ATCTTCTTCCACATTCGGAATAAAACCAGCTTC 933

Qy        149 ValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPhe 167

Db        932 ATATTTTATCTCAAAAATATAGTGGTAGCTGACCTCATCATGACCCTGACATTCCCATTC 873

Qy        168 LysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAla 187

Db        872 CGAATAGTCCGTGATGCAGGATTCGGACCTTGGTACTTCGAGTTTATCCTCTGCAGATAC 813

Qy        188 ThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSer 207

Db        812 ACCTCAGTTTTGTTCTATGCAAACATGTATACATCTATCGTGTTTCTTGGGCTGATCAGT 753

Qy        208 IleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThr 227

Db        752 GTTGATCGGTATCTAAAGGTGGTAAAGCCCTTTGGTGACTCTCGCATGTACAGCATAACC 693

Qy        228 TyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhe 247

Db        692 TTCACCAAAGTTTTATCAGTTTGTGTTTGGGTGATCATGGCTATTCTGTCCTTGCCAAAC 633

Qy        248 PheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspVal 267

Db        632 ATCATACTGACTAATGGG-----------CAGCCA---------ACTAAGGAAAATATT 594

Qy        268 HisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhe 287

Db        593 CATGACTGCATGAAACTCAAAAGTCCGCTGGGAGCCAAGTGGCATATGGCTGTCACCTAT 534

Qy        288 PheGly---------PheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIle 304

Db        533 GTGGACAGCTGTTTGTTTGTGGCCGTGCTGGTGATCCTGATTGGATGCTACATAGCCATC 474

Qy        305 IleArgThrLeuAsnAlaTyrAspHisArgTrpLeu--------------------Trp 317

Db        473 TCCAGATACATCCACAAATCCAGCAGGCAATTCATAAGCCAATCGAGCCGGAAGCGAAAG 414

Qy        318 TyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsn 337
```

```
Db        413 CACAACCAGAGCATCCGCGTGGTCGTGGCTGTGTTTTTTACCTGCTTCCTCCCGTATCAC 354

Qy        338 IleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly--------- 354
Db        353 TTGTGCAGAATC-----------CCCTTTACCTTCAGTAACTTAGACAGGCTTCTGGAT 306

Qy        355 ----------------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db        305 GAATCAGCACATAAAATCCTCTACTATTGCAAAGAAATGACACTTTTCTTGTCTGCGTGC 246

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db        245 AACGTGTGCCTGGATCCGATAATTTATTTTTTCATGTGTAAGTCA 201

RESULT 24
US-10-225-567A-224
; Sequence 224, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 224
;   LENGTH: 2299
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-224

Alignment Scores:
Pred. No.:              3.48e-26        Length:         2299
Score:                  387.00          Matches:        104
Percent Similarity:     45.36%          Conservative:   62
Best Local Similarity:  28.42%          Mismatches:     138
Query Match:            17.94%          Indels:         62
DB:                     9               Gaps:           12

US-09-208-629F-6 (1-408) x US-10-225-567A-224 (1-2299)

Qy         38 CysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLys 57
Db         16 TGTCAGAGTGGTGAACCCCTGCAGCCAGCAGGCCTCCTGAAAAAA--------------- 60

Qy         58 ThrPheArgGlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGly 77
Db         61 ------AAGTCCATGGGTGACAGAAGATTCATTGACTTCCAATTC--------------- 99

Qy         78 TrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeu 97
Db        100 ----------------------------------------CAAGATTCAAATTCAAGCCTC 120

Qy         98 HisValLys------AsnAlaThrMetGly---------TyrLeuThrSerSerLeuSer 112
Db        121 AGACCCAGGTTGGGCAATGCTACTGCCAATAATACTTGCATTGTTGATGATTCCTTCAAG 180

Qy        113 ThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAla 132
Db        181 TATAATCTCAATGGTGCTGTCTACAGTGTTGTATTCATCTTGGGTCTGATAACCAACAGT 240

Qy        133 ValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys---ThrThrValPheTyr 151
Db        241 GTCTCTCTGTTTGTCTTCTGTTTCCGCATGAAAATGAGAAGTGAGACTGCTATTTTTATC 300

Qy        152 ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyr 171
Db        301 ACCAATCTAGCTGTCTCTGATTTGCTTTTTGTCTGTACACTACCTTTTAAAATATTTTAC 360

Qy        172 HisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIle 191
Db        361 AACTTCAAC---CGCCACTGGCCTTTTGGTGACACCCTCTGCAAGATCTCTGGAACTGCA 417

Qy        192 PheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyr 211
Db        418 TTCCTTACCAACATCTATGGGAGCATGCTCTTTCTCACCTGTATTAGTGTGGATCGTTTC 477

Qy        212 LeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuVal 231
Db        478 CTGGCCATTGTCTATCCTTTTCGATCTCGTACTATTAGGACTAGGAGGAATTCTGCCATT 537

Qy        232 ThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLys 251
Db        538 GTGTGTGCTGGTGTCTGGATCCTAGTCCTCAGTGGCGGTATTTCAGCCTCTTTGTTT--- 594

Qy        252 GlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCys 271
Db        595 -----------------------TCCACCACTAATGTCAACAATGCAACCACCACCTGC 630

Qy        272 GluSerSerSerProPheGlnLeuTyr------TyrPheIleSerLeuAlaPhePhe--- 288
Db        631 TTTGAAGGCTTCTCCAAACGTGTCTGGAAGACTTATTTATCCAAGATCACAATATTTATT 690

Qy        289 ---------GlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305
Db        691 GAAGTTGTTGGGTTTATCATTCCTCTAATATTGAATGTCTCTTGCTCTTCTGTGGTGCTG 750

Qy        306 Arg-----------------ThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrVal 319
Db        751 AGAACTCTTCGCAAGCCTGCTACTCTGTCTCAAATTGGGACCAATAAGAAAAAAGTACTG 810

Qy        320 LysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIle 339
Db        811 AAAATGATCACAGTACATATGGCAGTCTTTGTGGTATGCTTTGTACCCTACAACTCTGTC 870

Qy        340 LeuIleIleHisHis--------------------AlaAsnTyrTyrTyrAsnAsnThr 352
Db        871 CTCTTCTTGTATGCCCTGGTGCGCTCCCAAGCTATTACTAATTGCTTTTTGGAAAGATTT 930

Qy        353 AspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 372
Db        931 GCAAAGATC------ATGTACCCAATCACCTTGTGCCTTGCAACTCTGAACTGTTGTTTT 984

Qy        373 AspProPheLeuTyrPhe 378
Db        985 GACCCTTTCATCTATTAC 1002

RESULT 25
US-09-077-173A-1
; Sequence 1, Application US/09077173A
; Publication No. US20030082674A1
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  RECEPTOR AND NUCLEIC ACID MOLECULE ENCODING
;    TITLE OF INVENTION:  SAID RECEPTOR
;    NUMBER OF SEQUENCES:  4
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/077,173A
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/BE 96/00123
;      FILING DATE:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1429 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  181..1275
US-09-077-173A-1

Alignment Scores:
Pred. No.:              1.98e-26        Length:         1429
Score:                  386.50          Matches:        121
Percent Similarity:     43.89%          Conservative:   55
Best Local Similarity:  30.17%          Mismatches:     166
Query Match:            17.92%          Indels:         60
DB:                     9               Gaps:           14

US-09-208-629F-6 (1-408) x US-09-077-173A-1 (1-1429)

Qy         30 LeuLeuLeuLeuLeuProThrPheCys---GlnSerGlyMetGluAsnAspThrAsnAsn 48
              ||||||   ||||||      ||||||      |||||||||                 |||
Db         80 CTGCTCTCCCTGCTCTACTTTTTTTGCTCCAGCTCAGGGATGGG-GGTGGGCAGGGAAAT 138

Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGlu 68
                 |||      |||   |||:::         |||||||||      ::::::
Db        139 CCTGCCACCCTCACTTCTCCCCTTCCCATCTCCAGGGGGGCCATGGCCAGTACAGAGTCC 198

Qy         69 GluPheProPheSer---AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
                        |||       |||            |||::::::      |||::::::
Db        199 TCCCTGTTGAGATCCCTAGGCCTCAGCCCAGGTCCTGGCAGCAGT---GAGGTGGAGCTG 255

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                |||                                                  :::
Db        256 GACTGT-------------------------------------------------TGGTTT 267

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
                                  |||:::|||      |||   :::|||||||||:::|||
Db        268 GATGAGGATTTCAAGTTCATCCTGCTGCCTGTGAGCTATGCAGTTGTCTTTGTGCTGGGC 327

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
                :::      ||||||   |||||||||:::      ||||||   |||
Db        328 TTGGGCCTTAACGCCCCAACCCTATGGCTCTTCATCTTCCGCCTCCGACCCTGGGATGCA 387

Qy        148 ThrValPheTyrThr---AsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                |||       |||      :::||||||:::::|||   |||:::   ::::::||||||
Db        388 ACGGCCACCTACATGTTCCACCTGGCATTGTCAGACACCTTGTATGTGCTGTCGCTGCCC 447

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
                   |||   |||:::         |||:::|||   ||||||      :::|||:::
Db        448 ACCCTCATCTACTATTATGCAGCCCACAACCACTGGCCCTTTGGCACTGAGATCTGCAAG 507

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                     :::||||||   |||:::|||||||||:::|||   |||   ||||||
Db        508 TTCGTCCGCTTTCTTTTCTATTGGAACCTCTACTGCAGTGTCCTTTTCCTCACCTGCATC 567

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
                |||::::::|||||||||   |||   ||||||                  :::
Db        568 AGCGTGCACCGCTACCTGGGCATCTGCCACCCACTTCGGGCACTACGCTGGGGCCGCCCT 627

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
                   |||   :::   |||      ||||||      |||         ::::::|||
Db        628 CGCCTCGCAGGCCTTCTCTGCCTGGCAGTTTGGTTGGTCGTAGCCGGCTGCCTCGTGCCC 687

Qy        247 ------PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCys 264
                    ||||||:::                 :::                   |||
Db        688 AACCTGTTCTTTGTCACAACCAGCAAC-----------AAAGGGACCACCGTCCTGTGC 735

Qy        265 HisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSer 284
                |||||   .      |||          |||      |||      |||
Db        736 CATGACACCACTCGGCCTGAAGAG-----------TTTGACCACTATGTGCACTTCAGC 783

Qy        285 LeuAlaPhePheGlyPheLeu------IleProPheValLeuIleIleTyrCysTyrAla 302
                  |||      |||   |||      :::|||   ::::::   :::   ||||||
```

```
Db        784 TCGGCGGTCATGGGGCTGCTCTTTGGCGTGCCCTGCCTGGTCACTCTTGTTTGCTATGGA 843

Qy        303 AlaIleIleArgThrLeu----------------AsnAlaTyrAspHisArgTrpLeu 316
                 :::   |||   |||                :::|||                |||
Db        844 CTCATGGCTCGTCGCCTGTATCAGCCCTTGCCAGGCTCTGCACAGTCGTCTTCTCGCCTC 903

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
                    ::::::         ::::::|||   :::|||   :::||||||   |||
Db        904 CGCTCTCTCCGCACCATAGCTGTGGTGCTGACTGTCTTTGCTGTCTGCTTCGTGCCTTTC 963

Qy        337 AsnIleIleLeuIleIleHisHis----------------AlaAsnTyrTyrTyrAsn 350
                 :::|||         |||::::::                    |||:::
Db        964 CACATCACCCGCACCATTTACTACCTGGCCAGGCTGTTGGAAGCTGACTGCCGAGTACTG 1023

Qy        351 AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSer 370
                 |||         :::      :::|||   :::          |||   |||   ||||||
Db       1024 AACATT------GTCAACGTGGTCTATAAAGTGACTCGGCCCCTGGCCAGTGCCAACAGC 1077

Qy        371 CysLeuAspProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyr 390
                 ||||||||||||   ||||||   |||
Db       1078 TGCCTGGATCCTGTGCTCTACTTGCTCACTGGG-------------------------- 1110

Qy        391 LeuThrLys******AsnAspLeuArgGlu---------GlnGlyGlnProSerGlnArg 407
                    |||           |||||:::           |||:::|||      |||
Db       1111 ---GACAAATATCGACGTCAGCTCCGTCAGCTCTGTGGTGGTGGCAAGCCCCAGCCCCGC 1167

Qy        408 Thr 408
                 |||
Db       1168 ACG 1170

RESULT 26
US-10-225-567A-220
; Sequence 220, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 220
;   LENGTH: 1118
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-220

Alignment Scores:
Pred. No.:              2.39e-26        Length:         1118
Score:                  384.00          Matches:        101
Percent Similarity:     48.04%          Conservative:   58
Best Local Similarity:  30.51%          Mismatches:     122
Query Match:            17.80%          Indels:         50
DB:                     9               Gaps:           11

US-09-208-629F-6 (1-408) x US-10-225-567A-220 (1-1118)

Qy         91 GluGluSerAlaSerHisLeuHisVal----------------------LysAsnAla 102
                 :::|||||||||||   |||   :::                      |||:::
Db          3 GATGAAAGTGCTTCCAAACTGAAAATTGGACGTGCCTTTACGATGGTAAGCGTTAACAGC 62

Qy        103 ThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeu 122
                 :::      :::          |||          |||       ::::::   :::
Db         63 TCCCACTGCTTCTATAATGACTCCTTTAAGTACACTTTGTATGGGTGCATGTTCAGCATG 122

Qy        123 ValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThr 142
                 |||||||||:::|||:::   :::|||   |||   :::::::::
```

```
Db        123 GTGTTTGTGCTTGGGTTAGTATCCAATTGTGTTGCCATATACATTTTC------------ 170

Qy        143 ArgSerIleCys--------------------ThrThrValPheTyrThrAsnLeuAla 155
                  ||||||                    ||||||   :::      |||||||||
Db        171 ------ATCTGCGTCCTCAAAGTCCGAAATGAAACTACAACTTACATGATTAACTTGGCA 224

Qy        156 IleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGly 175
              ::::::|||   ||||||       ||||||||||||:::|||   ::::::
Db        225 ATGTCAGACTTGCTTTTTGTTTTACTTTACCCTTCAGGATT---TTTTACTTCACAACA 281

Qy        176 AsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsn 195
                ||||||   ||||||::::::||||||:::   :::   ::::::||||||   |||
Db        282 CGGAATTGGCCATTTGGAGATTTACTTTGTAAGATTTCTGTGATGCTGTTTTATACCAAC 341

Qy        196 MetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleVal 215
              ||||||   |||||||||   |||   ||||||||||:::::::|||:::||||||||||||
Db        342 ATGTACGGAAGCATTCTGTTCTTAACCTGTATTAGTGTAGATCGATTTCTGGCAATTGTC 401

Qy        216 HisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu 235
              :::||||||      :::   |||          |||   :::   |||
Db        402 TACCCATTTAAGTCAAAGACTCTAAGAACCAAAAGAAATGCAAAGATTGTTTGCACTGGC 461

Qy        236 ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyr 255
              ||||||   ||||||              |||
Db        462 GTGTGGTTAACTGTGATCGGAGGAAGTGCACCC--------------------GCCGTT 500

Qy        256 LeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSer 275
                ||||||         :::      ::::::          |||      :::
Db        501 TTTGTTCAGTCTACCCACTCTCAGGGTAACAATGCCTCAGAAGCCTGCTTTGAAAATTTT 560

Qy        276 Pro-----------------------PheGlnLeuTyrTyrPheIleSerLeuAlaPhe 287
              |||                      ::::::      ||||||   :::
Db        561 CCAGAAGCCACATGGAAAACATATCTCTCAAGGATTGTAATTTTCATCGAAATAGTG--- 617

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
              ||||||   ||||||   :::|||   :::   |||   :::   :::::::::|||
Db        618 ---GGATTTTTTATTCCTCTAATTTTAAATGTAACTTGTTCTAGTATGGTGCTAAAAACT 674

Qy        308 LeuAsn---------AlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeu 324
              |||                      :::          |||      :::
Db        675 TTAACCAAACCAGTTACATTAAGTAGAAGCAAAATAAACAAAACTAAGGTTTTAAAAATG 734

Qy        325 Ile---------LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
              |||         |||:::||||||      ||||||   |||  ||||||   ||||||
Db        735 ATTTTTGTACATTTGATCATATTCTGTTTCTGTTTTGTTCCTTACAATATCAATCTTATT 794

Qy        342 IleHis---------HisAlaAsnTyrTyrTyrAsnAsnThr------AspGlyLeuTyr 356
              ::::::                  :::   |||   :::           :::
Db        795 TTATATTCTCTTGTGAGAACACAAACATTTGTTAATTGCTCAGTAGTGGCAGCAGTAAGG 854

Qy        357 PheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu 376
              :::|||   |||   |||||:::          |||   |||   ||||||   :::
Db        855 ACAATGTACCCAATCACTCTCTGTATTGCTGTTTCCAACTGTTGTTTTGACCCTATAGTT 914

Qy        377 TyrPheLeuMetSerLysThrArgAsnHisSer 387
              |||:::      |||   |||      :::|||
Db        915 TACTACTTTACATCGGACACAATTCAGAATTCA 947

RESULT 27
US-10-225-567A-331
; Sequence 331, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 331
;   LENGTH: 1098
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-331

Alignment Scores:
Pred. No.:                3.57e-26        Length:          1098
Score:                    382.00          Matches:         100
Percent Similarity:       46.82%          Conservative:    47
Best Local Similarity:    31.85%          Mismatches:      127
Query Match:              17.71%          Indels:          40
DB:                       9               Gaps:            10

US-09-208-629F-6 (1-408) x US-10-225-567A-331 (1-1098)

Qy        115 LeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThr 134
              |||:::|||      |||   :::|||||||||:::|||:::      ||||||   |||
Db        109 CTGCTGCCTGTGAGCTATGCAGTTGTCTTTGTGCTGGGCTTGGGCCTTAACGCCCCAACC 168

Qy        135 LeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThr---Asn 153
              ||||||:::      ||||||   |||              |||      |||      :::
Db        169 CTATGGCTCTTCATCTTCCGCCTCCGACCCTGGGATGCAACGGCCACCTACATGTTCCAC 228

Qy        154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
              ||||||:::::|||   |||:::   ::::::||||||      |||  |||:::
Db        229 CTGGCATTGTCAGACACCTTGTATGTGCTGTCGCTGCCCACCCTCATCTACTATTATGCA 288

Qy        174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
                  |||:::|||   ||||||      :::|||:::            :::||||||
Db        289 GCCCACAACCACTGGCCCTTTGGCACTGAGATCTGCAAGTTCGTCCGCTTTCTTTTCTAT 348

Qy        194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
                ||:::|||||||||:::|||   |||   |||||||||:::::|||||||||
Db        349 TGGAACCTCTACTGCAGTGTCCTTTTCCTCACCTGCATCAGCGTGCACCGCTACCTGGGC 408

Qy        214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233
              |||   ||||||              :::          |||   :::   |||
Db        409 ATCTGCCACCCACTTCGGGCACTACGCTGGGGCCGCCCTCGCCTCGCAGGCCTTCTCTGC 468

Qy        234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro------PhePheIleLeuLys 251
                  ||||||      |||          ::::::|||      |||||||:::
Db        469 CTGGCAGTTTGGTTGGTCGTAGCCGGCTGCCTCGTGCCCAACCTGTTCTTTGTCACAACC 528

Qy        252 GlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCys 271
                        :::                  |||||||||
Db        529 AGCAAC------------AAAGGGACCACCGTCCTGTGCCATGACACCACTCGGCCTGAA 576

Qy        272 GluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeu 291
              |||                  |||      |||      |||   |||      |||   |||
Db        577 GAG------------TTTGACCACTATGTGCACTTCAGCTCGGCGGTCATGGGGCTGCTC 624

Qy        292 ------IleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeu--- 308
                    :::|||   ::::::   :::   ||||||      :::   |||   |||
Db        625 TTTGGCGTGCCCTGCCTGGTCACTCTTGTTTGCTATGGACTCATGGCTCGTCGCCTGTAT 684

Qy        309 --------------AsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeu 323
                           :::|||                |||      ::::::
Db        685 CAGCCCTTGCCAGGCTCTGCACAGTCGTCTTCTCGCCTCCGCTCTCTCCGCACCATAGCT 744

Qy        324 LeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHis 343
              :::::::|||   :::|||   :::||||||   |||   :::|||      |||:::
Db        745 GTGGTGCTGACTGTCTTTGCTGTCTGCTTCGTGCCTTTCCACATCACCCGCACCATTTAC 804

Qy        344 His-----------------AlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPhe 357
              :::                 |||:::            |||      :::
Db        805 TACCTGGCCAGGCTGTTGGAAGCTGACTGCCGAGTACTGAACATT------GTCAACGTG 858

Qy        358 IleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyr 377
              :::|||   :::         |||   |||   |||||||||||||||||   ||||||
Db        859 GTCTATAAAGTGACTCGGCCCCTGGCCAGTGCCAACAGCTGCCTGGATCCTGTGCTCTAC 918
```

```
Qy        378 PheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAsp 397
Db        919 TTGCTCACTGGG----------------------------GACAAATATCGACGTCAG 948

Qy        398 LeuArgGlu---------GlnGlyGlnProSerGlnArgThr 408
Db        949 CTCCGTCAGCTCTGTGGTGGTGGCAAGCCCCAGCCCCGCACG 990
```

RESULT 28
US-09-828-478-1
; Sequence 1, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-828-478-1

Alignment Scores:

Pred. No.:	1.07e-25	Length:	1041
Score:	376.50	Matches:	93
Percent Similarity:	50.16%	Conservative:	62
Best Local Similarity:	30.10%	Mismatches:	129
Query Match:	17.45%	Indels:	25
DB:	9	Gaps:	9

US-09-208-629F-6 (1-408) x US-09-828-478-1 (1-1041)

```
Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db         31 TCCATCTCCGTATCAGAAATGGAACCAAATGGCACCTTCAGCAATAACAACAGCAGGAAC 90

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db         91 TGCACAATT-----------GAAAACTTCAAGAGAGAATTTTTCCCAATTGTATATCTG 138

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db        139 ATAATATTTTTCTGGGGAGTCTTGGGAAATGGGTTGTCCATATAT---GTTTTCCTGCAG 195

Qy        142 ------ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db        196 CCTTATAAGAAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGGCCATTTCAGATCTC 255

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db        256 CTGTTCATAAGCACGCTTCCCTTCAGGGCTGACTATTATCTTAGAGGCTCCAATTGGATA 315

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db        316 TTTGGAGACCTGGCCTGCAGGATTATGTCTTATTCCTTGTATGTCAACATGTACAGCAGT 375

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db        376 ATTTATTTCCTGACCGTGCTGAGTGTTGTGCGTTTCCTGGCAATGGTTCACCCCTTTCGG 435

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db        436 CTTCTGCATGTCACCAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG------ 489

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db        490 ---ATCCTTATCATGGCTTCCTCAATAATGCTCCTGGACAGTGGCTCTGAGCAGAACGGC 546

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
Db        547 AGTGTCACATCATGCTTAGAGCTGAAT------CTCTATAAAATTGCTAAGCTGCAGACC 600

Qy        280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299
Db        601 ATGAACTATATTGCCTTGGTG---GTGGGCTGCCTGCTGCCATTTTTCACACTCAGCATC 657

Qy        300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp---------- 312
Db        658 TGTTATCTGCTGATCATTCGGGTTCTGTTAAAAGTGGAGGTCCCAGAATCGGGGCTGCGG 717

Qy        313 ------HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThr 330
Db        718 GTTTCTCACAGGAAGGCA------CTGACCACCATCATCATCACCTTGATCATCTTCTTC 771

Qy        331 IleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsn 350
Db        772 TTGTGTTTCCTGCCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGT 831

Qy        351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
Db        832 TTATGCAAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCCAAT 891

Qy        370 SerCysLeuAspProPheLeuTyrPhe 378
Db        892 GCCTGCTTCAATCCTCTGCTCTATTAC 918
```

RESULT 29
US-09-826-791-5
; Sequence 5, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: PC10914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/198,367
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1041)
US-09-826-791-5

Alignment Scores:

Pred. No.:	1.07e-25	Length:	1041
Score:	376.50	Matches:	93
Percent Similarity:	50.16%	Conservative:	62
Best Local Similarity:	30.10%	Mismatches:	129
Query Match:	17.45%	Indels:	25
DB:	10	Gaps:	9

US-09-208-629F-6 (1-408) x US-09-826-791-5 (1-1041)

```
Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db         31 TCCATCTCCGTATCAGAAATGGAACCAAATGGCACCTTCAGCAATAACAACAGCAGGAAC 90

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
```

```
Db         91 TGCACAATT-----------GAAAACTTCAAGAGAGAATTTTTCCCAATTGTATATCTG 138

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141

Db        139 ATAATATTTTTCTGGGGAGTCTTGGGAAATGGGTTGTCCATATAT---GTTTTCCTGCAG 195

Qy        142 ------ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159

Db        196 CCTTATAAGAAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGGCCATTTCAGATCTC 255

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179

Db        256 CTGTTCATAAGCACGCTTCCCTTCAGGGCTGACTATTATCTTAGAGGCTCCAATTGGATA 315

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199

Db        316 TTTGGAGACCTGGCCTGCAGGATTATGTCTTATTCCTTGTATGTCAACATGTACAGCAGT 375

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219

Db        376 ATTTATTTCCTGACCGTGCTGAGTGTTGTGCGTTTCCTGGCAATGGTTCACCCCTTTCGG 435

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239

Db        436 CTTCTGCATGTCACCAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG------ 489

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259

Db        490 ---ATCCTTATCATGGCTTCCTCAATAATGCTCCTGGACAGTGGCTCTGAGCAGAACGGC 546

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279

Db        547 AGTGTCACATCATGCTTAGAGCTGAAT------CTCTATAAAATTGCTAAGCTGCAGACC 600

Qy        280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299

Db        601 ATGAACTATATTGCCTTGGTG---GTGGGCTGCCTGCTGCCATTTTTCACACTCAGCATC 657

Qy        300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp---------------------- 312

Db        658 TGTTATCTGCTGATCATTCGGGTTCTGTTAAAAGTGGAGGTCCCAGAATCGGGGCTGCGG 717

Qy        313 ------HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThr 330

Db        718 GTTTCTCACAGGAAGGCA------CTGACCACCATCATCATCACCTTGATCATCTTCTTC 771

Qy        331 IleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsn 350

Db        772 TTGTGTTTCCTGCCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGT 831

Qy        351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369

Db        832 TTATGCAAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCCAAT 891

Qy        370 SerCysLeuAspProPheLeuTyrPhe 378

Db        892 GCCTGCTTCAATCCTCTGCTCTATTAC 918

RESULT 30
US-09-866-230-6
; Sequence 6, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
;  APPLICANT: Murphy, Andrew, et al.
;  TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
;  FILE REFERENCE: REG 771A
;  CURRENT APPLICATION NUMBER: US/09/866,230
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: 60/207,725
;  PRIOR FILING DATE: 2000-05-26
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6
;   LENGTH: 1041
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1041)
US-09-866-230-6

Alignment Scores:
Pred. No.:                1.07e-25      Length:          1041
Score:                    376.50        Matches:         93
Percent Similarity:       50.16%        Conservative:    62
Best Local Similarity:    30.10%        Mismatches:      129
Query Match:              17.45%        Indels:          25
DB:                       10            Gaps:            9

US-09-208-629F-6 (1-408) x US-09-866-230-6 (1-1041)

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101

Db         31 TCCATCTCCGTATCAGAAATGGAACCAAATGGCACCTTCAGCAATAACAACAGCAGGAAC 90

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121

Db         91 TGCACAATT-----------GAAAACTTCAAGAGAGAATTTTTCCCAATTGTATATCTG 138

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141

Db        139 ATAATATTTTTCTGGGGAGTCTTGGGAAATGGGTTGTCCATATAT---GTTTTCCTGCAG 195

Qy        142 ------ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159

Db        196 CCTTATAAGAAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGGCCATTTCAGATCTC 255

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179

Db        256 CTGTTCATAAGCACGCTTCCCTTCAGGGCTGACTATTATCTTAGAGGCTCCAATTGGATA 315

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199

Db        316 TTTGGAGACCTGGCCTGCAGGATTATGTCTTATTCCTTGTATGTCAACATGTACAGCAGT 375

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219

Db        376 ATTTATTTCCTGACCGTGCTGAGTGTTGTGCGTTTCCTGGCAATGGTTCACCCCTTTCGG 435

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239

Db        436 CTTCTGCATGTCACCAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG------ 489

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259

Db        490 ---ATCCTTATCATGGCTTCCTCAATAATGCTCCTGGACAGTGGCTCTGAGCAGAACGGC 546

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279

Db        547 AGTGTCACATCATGCTTAGAGCTGAAT------CTCTATAAAATTGCTAAGCTGCAGACC 600

Qy        280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299

Db        601 ATGAACTATATTGCCTTGGTG---GTGGGCTGCCTGCTGCCATTTTTCACACTCAGCATC 657

Qy        300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp---------------------- 312

Db        658 TGTTATCTGCTGATCATTCGGGTTCTGTTAAAAGTGGAGGTCCCAGAATCGGGGCTGCGG 717

Qy        313 -----HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThr 330

Db        718 GTTTCTCACAGGAAGGCA------CTGACCACCATCATCATCACCTTGATCATCTTCTTC 771

Qy        331 IleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsn 350
```

```
Db        772 TTGTGTTTCCTGCCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGT 831

Qy        351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
                  |||   |||:::         :::|||   |||   |||   :::   |||
Db        832 TTATGCAAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCCAAT 891

Qy        370 SerCysLeuAspProPheLeuTyrPhe 378
              :::|||   :::|||   ||||||:::
Db        892 GCCTGCTTCAATCCTCTGCTCTATTAC 918

RESULT 31
US-09-779-679-1
; Sequence 1, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond J
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Mishra, Vishnu S
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Ballinger, Robert
;  APPLICANT:  Vernet, Corine A
;  APPLICANT:  Li, Li
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Andrew, David P
;  APPLICANT:  Mezes, Peter S
;  TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
;  FILE REFERENCE: 15966-661
;  CURRENT APPLICATION NUMBER: US/09/779,679
;  CURRENT FILING DATE:  2001-02-08
;  PRIOR APPLICATION NUMBER: USSN 60/181045
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: USSN 60/183191
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: USSN 60/180,929
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: USSN 60/219758
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: USSN 60/181339
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: USSN 60/181344
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: USSN 60/221341
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: USSN 60/181392
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: USSN 60/219585
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: USSN 60/181157
;  PRIOR FILING DATE: 2000-02-09
;  NUMBER OF SEQ ID NOS: 65
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1260
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (105)..(1142)
;   NAME/KEY: misc_feature
;   LOCATION: (120)..(122)
;   OTHER INFORMATION: Alternative ATG start site
US-09-779-679-1

Alignment Scores:
Pred. No.:              1.4e-25       Length:          1260
Score:                  376.50        Matches:         93
Percent Similarity:     50.16%        Conservative:    62
Best Local Similarity:  30.10%        Mismatches:      129
Query Match:            17.45%        Indels:          25
DB:                     9             Gaps:            9

US-09-208-629F-6 (1-408) x US-09-779-679-1 (1-1260)

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
              :::|||:::|||           |||      :::   |||:::   :::   :::|||
Db        135 TCCATCTCCGTATCAGAAATGGAACCAAATGGCACCTTCAGCAATAACAACAGCAGGAAC 194

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
                 |||:::            :::      :::      |||   :::||||||
Db        195 TGCACAATT-----------GAAAACTTCAAGAGAGAATTTTTCCCAATTGTATATCTG 242

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              ::::::|||      ||||||      |||   ::::::::::::   :::|||   :::
Db        243 ATAATATTTTTCTGGGGAGTCTTGGGAAATGGGTTGTCCATATAT---GTTTTCCTGCAG 299

Qy        142 ------ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
                    :::|||           ||||||      |||||||||||||:::|||
Db        300 CCTTATAAGAAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGGCCATTTCAGATCTC 359

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
              ||||||      |||||||||||:::      |||:::|||   |||:::||||||:::
Db        360 CTGTTCATAAGCACGCTTCCCTTCAGGGCTGACTATTATCTTAGAGGCTCCAATTGGATA 419

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
              ||||||::::::  ||||||      :::           |||   |||||||||   |||
Db        420 TTTGGAGACCTGGCCTGCAGGATTATGTCTTATTCCTTGTATGTCAACATGTACAGCAGT 479

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
              |||      |||      :::|||:::   |||:::||||||:::||||||||||||
Db        480 ATTTATTTCCTGACCGTGCTGAGTGTTGTGCGTTTCCTGGCAATGGTTCACCCCTTTCGG 539

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
                       :::           |||   :::   ||||||::::::|||
Db        540 CTTCTGCATGTCACCAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG------ 593

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
                    |||   |||           :::|||
Db        594 ---ATCCTTATCATGGCTTCCTCAATAATGCTCCTGGACAGTGGCTCTGAGCAGAACGGC 650

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
              :::|||:::|||   :::::::::                  :::      |||
Db        651 AGTGTCACATCATGCTTAGAGCTGAAT------CTCTATAAAATTGCTAAGCTGCAGACC 704

Qy        280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299
                 :::|||:::|||           |||   |||:::||||||      :::
Db        705 ATGAACTATATTGCCTTGGTG---GTGGGCTGCCTGCTGCCATTTTTCACACTCAGCATC 761

Qy        300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp---------------------- 312
              ||||||      ||||||||   |||          :::
Db        762 TGTTATCTGCTGATCATTCGGGTTCTGTTAAAAGTGGAGGTCCCAGAATCGGGGCTGCGG 821

Qy        313 ------HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThr 330
                    ||||||           :::          ::::::   |||:::||||||
Db        822 GTTTCTCACAGGAAGGCA------CTGACCACCATCATCATCACCTTGATCATCTTCTTC 875

Qy        331 IleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsn 350
              :::||||||   |||   :::   :::      :::|||          :::
Db        876 TTGTGTTTCCTGCCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGT 935

Qy        351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
                       |||   |||:::         :::|||   |||   |||   :::   |||
Db        936 TTATGCAAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCCAAT 995

Qy        370 SerCysLeuAspProPheLeuTyrPhe 378
              :::|||   :::|||   ||||||:::
Db        996 GCCTGCTTCAATCCTCTGCTCTATTAC 1022

RESULT 32
US-09-779-679-24
; Sequence 24, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond J
;  APPLICANT:  Burgess, Catherine E
```

```
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Mishra, Vishnu S
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Ballinger, Robert
;  APPLICANT:  Vernet, Corine A
;  APPLICANT:  Li, Li
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Andrew, David P
;  APPLICANT:  Mezes, Peter S
;  TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
;  FILE REFERENCE: 15966-661
;  CURRENT APPLICATION NUMBER: US/09/779,679
;  CURRENT FILING DATE:  2001-02-08
;  PRIOR APPLICATION NUMBER: USSN 60/181045
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: USSN 60/183191
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: USSN 60/180,929
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: USSN 60/219758
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: USSN 60/181339
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: USSN 60/181344
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: USSN 60/221341
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: USSN 60/181392
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: USSN 60/219585
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: USSN 60/181157
;  PRIOR FILING DATE: 2000-02-09
;  NUMBER OF SEQ ID NOS: 65
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 1260
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-779-679-24

Alignment Scores:
Pred. No.:                1.4e-25        Length:          1260
Score:                    376.50         Matches:         93
Percent Similarity:       50.16%         Conservative:    62
Best Local Similarity:    30.10%         Mismatches:      129
Query Match:              17.45%         Indels:          25
DB:                       9              Gaps:            9

US-09-208-629F-6 (1-408) x US-09-779-679-24 (1-1260)

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
              :::|||:::|||             |||      :::   |||:::   :::   :::|||
Db        135 TCCATCTCCGTATCAGAAATGGAACCAAATGGCACCTTCAGCAATAACAACAGCAGGAAC 194

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
              |||:::          :::         :::      |||   :::||||||
Db        195 TGCACAATT-----------GAAAACTTCAAGAGAGAATTTTTCCCAATTGTATATCTG 242

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
              ::::::|||      ||||||      |||   ::::::::::::   :::|||   :::
Db        243 ATAATATTTTTCTGGGGAGTCTTGGGAAATGGGTTGTCCATATAT---GTTTTCCTGCAG 299

Qy        142 ------ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
                  :::|||             ||||||      |||||||||||:::|||
Db        300 CCTTATAAGAAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGGCCATTTCAGATCTC 359

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
              ||||||      ||||||||||||:::      |||:::|||   |||:::||||||:::
Db        360 CTGTTCATAAGCACGCTTCCCTTCAGGGCTGACTATTATCTTAGAGGCTCCAATTGGATA 419

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
              ||||||::::::   ||||||      :::         |||   |||||||||   |||
Db        420 TTTGGAGACCTGGCCTGCAGGATTATGTCTTATTCCTTGTATGTCAACATGTACAGCAGT 479

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
              |||      |||      :::|||:::   |||:::||||||:::||||||||||||
Db        480 ATTTATTTCCTGACCGTGCTGAGTGTTGTGCGTTTCCTGGCAATGGTTCACCCCTTTCGG 539

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
                  :::                |||  :::   ||||||::::::|||
Db        540 CTTCTGCATGTCACCAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG------ 593

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
              |||   |||            :::|||
Db        594 ---ATCCTTATCATGGCTTCCTCAATAATGCTCCTGGACAGTGGCTCTGAGCAGAACGGC 650

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
              :::|||:::|||   :::::::::                  :::      |||
Db        651 AGTGTCACATCATGCTTAGAGCTGAAT------CTCTATAAAATTGCTAAGCTGCAGACC 704

Qy        280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299
              :::|||:::|||         |||  |||:::||||||      :::
Db        705 ATGAACTATATTGCCTTGGTG---GTGGGCTGCCTGCTGCCATTTTTCACACTCAGCATC 761

Qy        300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp---------- 312
              ||||||      ||||||||  |||      :::
Db        762 TGTTATCTGCTGATCATTCGGGTTCTGTTAAAAGTGGAGGTCCCAGAATCGGGGCTGCGG 821

Qy        313 ------HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThr 330
                  ||||||          :::         ::::::   |||:::||||||
Db        822 GTTTCTCACAGGAAGGCA------CTGACCACCATCATCATCACCTTGATCATCTTCTTC 875

Qy        331 IleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsn 350
              :::||||||   |||   :::   :::      :::|||         :::
Db        876 TTGTGTTTCCTGCCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGT 935

Qy        351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
                    |||   |||:::      :::|||   |||   |||   :::   |||
Db        936 TTATGCAAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCCAAT 995

Qy        370 SerCysLeuAspProPheLeuTyrPhe 378
              :::|||   :::|||   ||||||:::
Db        996 GCCTGCTTCAATCCTCTGCTCTATTAC 1022

RESULT 33
US-09-828-478-3
; Sequence 3, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
;  APPLICANT: Xiao, Yonghong
;  TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
;  TITLE OF INVENTION:  Protein
;  FILE REFERENCE: 04974.00458
;  CURRENT APPLICATION NUMBER: US/09/828,478
;  CURRENT FILING DATE:  2001-04-09
;  PRIOR APPLICATION NUMBER: 60/195,196
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: 60/254,876
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 1430
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-828-478-3

Alignment Scores:
Pred. No.:                1.68e-25       Length:          1430
Score:                    376.50         Matches:         93
Percent Similarity:       50.16%         Conservative:    62
Best Local Similarity:    30.10%         Mismatches:      129
Query Match:              17.45%         Indels:          25
DB:                       9              Gaps:            9
```

```
US-09-208-629F-6 (1-408) x US-09-828-478-3 (1-1430)

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db        283 TCCATCTCCGTATCAGAAATGGAACCAAATGGCACCTTCAGCAATAACAACAGCAGGAAC 342

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db        343 TGCACAATT-----------GAAAACTTCAAGAGAGAATTTTTCCCAATTGTATATCTG 390

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db        391 ATAATATTTTTCTGGGGAGTCTTGGGAAATGGGTTGTCCATATAT---GTTTTCCTGCAG 447

Qy        142 ------ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db        448 CCTTATAAGAAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGGCCATTTCAGATCTC 507

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db        508 CTGTTCATAAGCACGCTTCCCTTCAGGGCTGACTATTATCTTAGAGGCTCCAATTGGATA 567

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db        568 TTTGGAGACCTGGCCTGCAGGATTATGTCTTATTCCTTGTATGTCAACATGTACAGCAGT 627

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db        628 ATTTATTTCCTGACCGTGCTGAGTGTTGTGCGTTTCCTGGCAATGGTTCACCCCTTTCGG 687

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db        688 CTTCTGCATGTCACCAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG------ 741

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db        742 ---ATCCTTATCATGGCTTCCTCAATAATGCTCCTGGACAGTGGCTCTGAGCAGAACGGC 798

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
Db        799 AGTGTCACATCATGCTTAGAGCTGAAT------CTCTATAAAATTGCTAAGCTGCAGACC 852

Qy        280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299
Db        853 ATGAACTATATTGCCTTGGTG---GTGGGCTGCCTGCTGCCATTTTTCACACTCAGCATC 909

Qy        300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp---------------------- 312
Db        910 TGTTATCTGCTGATCATTCGGGTTCTGTTAAAAGTGGAGGTCCCAGAATCGGGGCTGCGG 969

Qy        313 ------HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThr 330
Db        970 GTTTCTCACAGGAAGGCA------CTGACCACCATCATCATCACCTTGATCATCTTCTTC 1023

Qy        331 IleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsn 350
Db       1024 TTGTGTTTCCTGCCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGT 1083

Qy        351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
Db       1084 TTATGCAAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCCAAT 1143

Qy        370 SerCysLeuAspProPheLeuTyrPhe 378
Db       1144 GCCTGCTTCAATCCTCTGCTCTATTAC 1170

RESULT 34
US-09-728-952-26
; Sequence 26, Application US/09728952
; Patent No. US20020111302A1
; GENERAL INFORMATION:
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Vicki
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION:  No. US20020111302A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE:  799
;  CURRENT APPLICATION NUMBER: US/09/728,952
;  CURRENT FILING DATE:  2000-11-30
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  pt_FL_genes Version 2.0
; SEQ ID NO 26
;   LENGTH: 1700
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (232)..(1272)
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1700)
;   OTHER INFORMATION: n = a,t,c or g
US-09-728-952-26

Alignment Scores:
Pred. No.:                2.14e-25        Length:          1700
Score:                    376.50          Matches:         93
Percent Similarity:       50.16%          Conservative:    62
Best Local Similarity:    30.10%          Mismatches:      129
Query Match:              17.45%          Indels:          25
DB:                       10              Gaps:            9

US-09-208-629F-6 (1-408) x US-09-728-952-26 (1-1700)

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db        262 TCCATCTCCGTATCAGAAATGGAACCAAATGGCACCTTCAGCAATAACAACAGCAGGAAC 321

Qy        102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db        322 TGCACAATT-----------GAAAACTTCAAGAGAGAATTTTTCCCAATTGTATATCTG 369

Qy        122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db        370 ATAATATTTTTCTGGGGAGTCTTGGGAAATGGGTTGTCCATATAT---GTTTTCCTGCAG 426

Qy        142 ------ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db        427 CCTTATAAGAAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGGCCATTTCAGATCTC 486

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db        487 CTGTTCATAAGCACGCTTCCCTTCAGGGCTGACTATTATCTTAGAGGCTCCAATTGGATA 546

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db        547 TTTGGAGACCTGGCCTGCAGGATTATGTCTTATTCCTTGTATGTCAACATGTACAGCAGT 606

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db        607 ATTTATTTCCTGACCGTGCTGAGTGTTGTGCGTTTCCTGGCAATGGTTCACCCCTTTCGG 666

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db        667 CTTCTGCATGTCACCAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG------ 720

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db        721 ---ATCCTTATCATGGCTTCCTCAATAATGCTCCTGGACAGTGGCTCTGAGCAGAACGGC 777
```

```
Qy       260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
Db       778 AGTGTCACATCATGCTTAGAGCTGAAT------CTCTATAAAATTGCTAAGCTGCAGACC 831

Qy       280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299
Db       832 ATGAACTATATTGCCTTGGTG---GTGGGCTGCCTGCTGCCATTTTTCACACTCAGCATC 888

Qy       300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp---------------------- 312
Db       889 TGTTATCTGCTGATCATTCGGGTTCTGTTAAAAGTGGAGGTCCCAGAATCGGGGCTGCGG 948

Qy       313 ------HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThr 330
Db       949 GTTTCTCACAGGAAGGCA------CTGACCACCATCATCATCACCTTGATCATCTTCTTC 1002

Qy       331 IleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsn 350
Db      1003 TTGTGTTTCCTGCCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGT 1062

Qy       351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
Db      1063 TTATGCAAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCCAAT 1122

Qy       370 SerCysLeuAspProPheLeuTyrPhe 378
Db      1123 GCCTGCTTCAATCCTCTGCTCTATTAC 1149

RESULT 35
US-10-225-567A-588
; Sequence 588, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 588
;   LENGTH: 2807
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-588

Alignment Scores:
Pred. No.:              4.33e-25        Length:         2807
Score:                  376.50          Matches:        93
Percent Similarity:     50.16%          Conservative:   62
Best Local Similarity:  30.10%          Mismatches:     129
Query Match:            17.45%          Indels:         25
DB:                     9               Gaps:           9

US-09-208-629F-6 (1-408) x US-10-225-567A-588 (1-2807)

Qy        82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db       294 TCCATCTCCGTATCAGAAATGGAACCAAATGGCACCTTCAGCAATAACAACAGCAGGAAC 353

Qy       102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db       354 TGCACAATT-----------GAAAACTTCAAGAGAGAATTTTTCCCAATTGTATATCTG 401

Qy       122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db       402 ATAATATTTTTCTGGGGAGTCTTGGGAAATGGGTTGTCCATATAT---GTTTTCCTGCAG 458

Qy       142 ------ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db       459 CCTTATAAGAAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGGCCATTTCAGATCTC 518

Qy       160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db       519 CTGTTCATAAGCACGCTTCCCTTCAGGGCTGACTATTATCTTAGAGGCTCCAATTGGATA 578

Qy       180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db       579 TTTGGAGACCTGGCCTGCAGGATTATGTCTTATTCCTTGTATGTCAACATGTACAGCAGT 638

Qy       200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db       639 ATTTATTTCCTGACCGTGCTGAGTGTTGTGCGTTTCCTGGCAATGGTTCACCCCTTTCGG 698

Qy       220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db       699 CTTCTGCATGTCACCAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG------ 752

Qy       240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db       753 ---ATCCTTATCATGGCTTCCTCAATAATGCTCCTGGACAGTGGCTCTGAGCAGAACGGC 809

Qy       260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
Db       810 AGTGTCACATCATGCTTAGAGCTGAAT------CTCTATAAAATTGCTAAGCTGCAGACC 863

Qy       280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299
Db       864 ATGAACTATATTGCCTTGGTG---GTGGGCTGCCTGCTGCCATTTTTCACACTCAGCATC 920

Qy       300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp---------------------- 312
Db       921 TGTTATCTGCTGATCATTCGGGTTCTGTTAAAAGTGGAGGTCCCAGAATCGGGGCTGCGG 980

Qy       313 ------HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThr 330
Db       981 GTTTCTCACAGGAAGGCA------CTGACCACCATCATCATCACCTTGATCATCTTCTTC 1034

Qy       331 IleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsn 350
Db      1035 TTGTGTTTCCTGCCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGT 1094

Qy       351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
Db      1095 TTATGCAAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCCAAT 1154

Qy       370 SerCysLeuAspProPheLeuTyrPhe 378
Db      1155 GCCTGCTTCAATCCTCTGCTCTATTAC 1181

RESULT 36
US-09-826-791-1
; Sequence 1, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
;  APPLICANT: Pfizer Inc
;  TITLE OF INVENTION: No. US20010039037A1el Polypeptide
;  FILE REFERENCE: PC10914ADAM
;  CURRENT APPLICATION NUMBER: US/09/826,791
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 0008504.3
;  PRIOR FILING DATE: 2000-04-05
;  PRIOR APPLICATION NUMBER: 60/198,367
;  PRIOR FILING DATE: 2000-04-19
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 993
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-09-826-791-1

Alignment Scores:

Pred. No.:	1.54e-25	Length:	993
Score:	374.50	Matches:	91
Percent Similarity:	50.17%	Conservative:	60
Best Local Similarity:	30.23%	Mismatches:	125
Query Match:	17.36%	Indels:	25
DB:	10	Gaps:	9

US-09-208-629F-6 (1-408) x US-09-826-791-1 (1-993)

```
Qy         90 ProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSer 109
Db          7 CCAAATGGCACCTTCAGCAATAACAACAGCAGGAACTGCACAATT------------GAA 54

Qy        110 SerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValPro 129
Db         55 AACTTCAAGAGAGAATTTTTCCCAATTGTATATCTGATAATATTTTTCTGGGGAGTCTTG 114

Qy        130 AlaAsnAlaValThrLeuTrpMetLeuPhePheArg------ThrArgSerIleCysThr 147
Db        115 GGAAATGGGTTGTCCATATAT---GTTTTCCTGCAGCCTTATAAGAAGTCCACATCTGTG 171

Qy        148 ThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPhe 167
Db        172 AACGTTTTCATGCTAAATCTGGCCATTTCAGATCTCCTGTTCATAAGCACGCTTCCCTTC 231

Qy        168 LysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAla 187
Db        232 AGGGCTGACTATTATCTTAGAGGCTCCAATTGGATATTTGGAGACCTGGCCTGCAGGATT 291

Qy        188 ThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSer 207
Db        292 ATGTCTTATTCCTTGTATGTCAACATGTACAGCAGTATTTATTTCCTGACCGTGCTGAGT 351

Qy        208 IleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThr 227
Db        352 GTTGTGCGTTTCCTGGCAATGGTTCACCCCTTTCGGCTTCTGCATGTCACCAGCATCAGG 411

Qy        228 TyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhe 247
Db        412 AGTGCCTGGATCCTCTGTGGGATCATATGG---------ATCCTTATCATGGCTTCCTCA 462

Qy        248 PheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspVal 267
Db        463 ATAATGCTCCTGGACAGTGGCTCTGAGCAGAACGGCAGTGTCACATCATGCTTAGAGCTG 522

Qy        268 HisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhe 287
Db        523 AAT------CTCTATAAAATTGCTAAGCTGCAGACCATGAACTATATTGCCTTGGTG--- 573

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
Db        574 GTGGGCTGCCTGCTGCCATTTTTCACACTCAGCATCTGTTATCTGCTGATCATTCGGGTT 633

Qy        308 LeuAsnAlaTyrAsp--------------------------HisArgTrpLeuTrpTyr 318
Db        634 CTGTTAAAAGTGGAGGTCCCAGAATCGGGGCTGCGGGTTTCTCACAGGAAGGCA------ 687

Qy        319 ValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIle 338
Db        688 CTGACCACCATCATCATCACCTTGATCATCTTCTTCTTGTGTTTCCTGCCCTATCACACA 747

Qy        339 IleLeuIleIleHisHisAlaAsnTyrTyrTyrAsn---AsnThrAspGlyLeuTyrPhe 357
Db        748 CTGAGGACCGTCCACTTGACGACATGGAAAGTGGGTTTATGCAAAGACAGACTGCATAAA 807

Qy        358 IleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyr 377
Db        808 GCTTTGGTTATCACACTGGCCTTGGCAGCAGCCAATGCCTGCTTCAATCCTCTGCTCTAT 867

Qy        378 Phe 378
```

```
Db        868 TAC 870
```

RESULT 37
US-10-225-567A-546
; Sequence 546, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 546
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-546

Alignment Scores:

Pred. No.:	2.42e-25	Length:	1014
Score:	372.50	Matches:	103
Percent Similarity:	47.51%	Conservative:	59
Best Local Similarity:	30.21%	Mismatches:	140
Query Match:	17.27%	Indels:	39
DB:	9	Gaps:	12

US-09-208-629F-6 (1-408) x US-10-225-567A-546 (1-1014)

```
Qy         92 GluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeu 111
Db          4 GATGAAACAGGAAATCTGACAGTATCTTCTGCCACATGCCATGACACTATTGATGACTTC 63

Qy        112 SerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsn 131
Db         64 CGCAATCAAGTGTATTCCACCTTGTACTCTATGATCTCTGTTGTAGGCTTCTTTGGCAAT 123

Qy        132 AlaValThrLeuTrpMetLeuPhe------PheArgThrArgSerIleCysThrThrVal 149
Db        124 GGCTTTGTGCTCTATGTCCTCATAAAAACCTATCACAAGAAGTCA---GCCTTCCAAGTA 180

Qy        150 PheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIle 169
Db        181 TACATGATTAATTTAGCAGTAGCAGATCTACTTTGTGTGTGCACACTGCCTCTCCGTGTG 240

Qy        170 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThr 189
Db        241 GTCTATTATGTTCACAAAGGCATTTGGCTCTTTGGTGACTTCTTGTGCCGCCTCAGCACC 300

Qy        190 ValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsn 209
Db        301 TATGCTTTGTATGTCAACCTCTATTGTAGCATCTTCTTTATGACAGCCATGAGCTTTTTC 360

Qy        210 ArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAla 229
Db        361 CGGTGCATTGCAATTGTTTTTCCAGTCCAGAACATTAATTTGGTTACACAGAAAAAAGCC 420

Qy        230 LeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIle 249
Db        421 AGGTTTGTGTGTGTAGGTATTTGGATTTTTGTGATTTTGACCAGTTCTCCATTTCTAATG 480

Qy        250 LeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsn 269
Db        481 GCCAAACCACAA------AAAGATGAGAAAAATAATACCAAGTGCTTTGAGCCCCCACAA 534

Qy        270 ThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGly 289
```

```
Db        535 GACAATCAAACTAAAAATCATGTTTTGGTCTTGCATTATGTGTCATTG---TTTGTTGGC 591

Qy        290 PheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeu--- 308
Db        592 TTTATCATCCCTTTTGTTATTATAATTGTCTGTTACACAATGATCATTTTGACCTTACTA 651

Qy        309 ------------------AsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSer 322
Db        652 AAAAAATCAATGAAAAAAAATCTGTCAAGTCATAAA--------------AAGGCTATA 696

Qy        323 LeuLeuIleLeuValIle---------PheThrIleCysPheAlaProSerAsnIleIle 339
Db        697 GGAATGATCATGGTCGTGACCGCTGCCTTTTTAGTCAGTTTCATGCCATATCATATTCAA 756

Qy        340 LeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThr---------AspGlyLeuTyr 356
Db        757 CGTACCATT---CACCTTCATTTTTTACACAATGAAACTAAACCCTGTGATTCTGTCCTT 813

Qy        357 PheIle---------TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
Db        814 AGAATGCAGAAGTCCGTGGTCATAACCTTGTCTCTGGCTGCATCCAATTGTTGCTTTGAC 873

Qy        374 ProPheLeuTyrPheLeu-----------------------MetSerLysThrArgAsn 385
Db        874 CCTCTCCTATATTTCTTTTCTGGGGGTAACTTTAGGAAAAGGCTGTCTACATTCAGAAAG 933

Qy        386 His---------SerThrAlaTyrLeuThrLys******AsnAspLeuArgGluGlnGly 402
Db        934 CATTCTTTGTCCAGCGTGACTTATGTACCCAGAAAGAAGGCCTCTTTGCCAGAAAAAGGA 993

Qy        403 Gln 403
Db        994 GAA 996

RESULT 38
US-10-167-192-4
; Sequence 4, Application US/10167192
; Publication No. US20030040052A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  TITLE OF INVENTION: No. US20030040052Alel G-Protein Coupled Receptors
;  FILE REFERENCE: 5800-11A
;  CURRENT APPLICATION NUMBER: US/10/167,192
;  CURRENT FILING DATE:  2002-06-11
;  PRIOR APPLICATION NUMBER: US/09/420,187
;  PRIOR FILING DATE: 1999-10-18
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 1358
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-167-192-4

Alignment Scores:
Pred. No.:              5.04e-25        Length:         1358
Score:                  371.00          Matches:        106
Percent Similarity:     46.41%          Conservative:   62
Best Local Similarity:  29.28%          Mismatches:     144
Query Match:            17.20%          Indels:         50
DB:                     9               Gaps:           13

US-09-208-629F-6 (1-408) x US-10-167-192-4 (1-1358)

Qy         82 ThrIleThrValLysIleLysCysProGluGluSerAlaSer----------------- 95
Db        170 ACATTTGTGAAGGTACTCCAGTGCCAGAAAGAGGCACAAAGCAGACATTCGTAGAGAAAC 229

Qy         96 --------------HisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSer 110
Db        230 ATGGATGAAACAGGAAATCTGACAGTATCTTCTGCCACATGCCATGACACTATTGATGAC 289

Qy        111 LeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAla 130
Db        290 TTCCGCAATCAAGTGTATTCCACCTTGTACTCTATGATCTCTGTTGTAGGCTTCTTTGGC 349

Qy        131 AsnAlaValThrLeuTrpMetLeuPhe------PheArgThrArgSerIleCysThrThr 148
Db        350 AATGGCTTTGTGCTCTATGTCCTCATAAAAACCTATCACAAGAAGTCA---GCCTTCCAA 406

Qy        149 ValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLys 168
Db        407 GTATACATGATTAATTTAGCAGTAGCAGATCTACTTTGTGTGTGCACACTGCCTCTCCGT 466

Qy        169 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThr 188
Db        467 GTGGTCTATTATGTTCACAAAGGCATTTGGCTCTTTGGTGACTTCTTGTGCCGCCTCAGC 526

Qy        189 ThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIle 208
Db        527 ACCTATGCTTTGTATGTCAACCTCTATTGTAGCATCTTCTTTATGACAGCCATGAGCTTT 586

Qy        209 AsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyr 228
Db        587 TTCCGGTGCATTGCAATTGTTTTTCCAGTCCAGAACATTAATTTGGTTACACAGAAAAAA 646

Qy        229 AlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePhe 248
Db        647 GCCAGGTTTGTGTGTGTAGGTATTTGGATTTTTGTGATTTTGACCAGTTCTCCATTTCTA 706

Qy        249 IleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHis 268
Db        707 ATGGCCAAACCACAA------AAAGATGAGAAAAATAATACCAAGTGCTTTGAGCCCCCA 760

Qy        269 AsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePhe 288
Db        761 CAAGACAATCAAACTAAAAATCATGTTTTGGTCTTGCATTATGTGTCATTGGTT---GTT 817

Qy        289 GlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeu 308
Db        818 GGCTTTATCATCCCTTTTGTTATTATAATTGTCTGTTACACAATGATCATTTTGACCTTA 877

Qy        309 --------------------AsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAla 321
Db        878 CTAAAAAAATCAATGAAAAAAAATCTGTCAAGTCATAAA--------------AAGGCT 922

Qy        322 SerLeuLeuIleLeuValIle---------PheThrIleCysPheAlaProSerAsnIle 338
Db        923 ATAGGAATGATCATGGTCGTGACCGCTGCCTTTTTAGTCAGTTTCATGCCATATCATATT 982

Qy        339 IleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThr---------AspGlyLeu 355
Db        983 CAACGTACCATT---CACCTTCATTTTTTACACAATGAAACTAAACCCTGTGATTCTGTC 1039

Qy        356 TyrPheIle---------TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 372
Db       1040 CTTAGAATGCAGAAGTCCGTGGTCATAACCTTGTCTCTGGCTGCATCCAATTGTTGCTTT 1099

Qy        373 AspProPheLeuTyrPheLeu-----------------------MetSerLysThrArg 384
Db       1100 GACCCTCTCCTATATTTCTTTTCTGGGGGTAACTTTAGGAAAAGGCTGTCTACATTTAGA 1159

Qy        385 AsnHis---------SerThrAlaTyrLeuThrLys******AsnAspLeuArgGluGln 401
Db       1160 AAGCATTCTTTGTCCAGCGTGACTTATGTACCCAGAAAGAAGGCCTCTTTGCCAGAAAAA 1219

Qy        402 GlyGln 403
Db       1220 GGAGAA 1225

RESULT 39
US-09-826-508-9
; Sequence 9, Application US/09826508
; Patent No. US20010025099A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Nabil Elshourbagy
;  APPLICANT:  Lisa Vawter
;  TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
;  TITLE OF INVENTION:  and Polynucleotides
;  FILE REFERENCE: GP-70744USB
;  CURRENT APPLICATION NUMBER: US/09/826,508
;  CURRENT FILING DATE:  2001-04-05
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 1077
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-826-508-9

Alignment Scores:
Pred. No.:              6.2e-25       Length:          1077
Score:                  368.50        Matches:         89
Percent Similarity:     50.33%        Conservative:    64
Best Local Similarity:  29.28%        Mismatches:      116
Query Match:            17.08%        Indels:          35
DB:                     10            Gaps:            9

US-09-208-629F-6 (1-408) x US-09-826-508-9 (1-1077)

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
              ||||||   |||         |||   :::         |||   :::::::|||   :::
Db         94 AAGAACACCACC---------CTTCACAATGAATTTGACACAATTGTCTTGCCGGTGCTT 144

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              ||||||::::::||||||      :::      |||   :::   :::|||   :::|||
Db        145 TATCTCATTATATTTGTGGCAAGCATCTTGCTGAATGGTTTAGCAGTGTGG---ATCTTC 201

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||      |||:::      :::   :::||||||      |||:::   :::||||||
Db        202 TTCCACATTAGGAATAAAACCAGCTTCATATTCTATCTCAAAAACATAGTGGTTGCAGAC 261

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
              :::      :::|||   ||||||:::|||  :::                    |||
Db        262 CTCATAATGACGCTGACATTTCCATTTCGAATAGTCCATGATGCAGGATTTGGACCTTGG 321

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              |||      :::|||||||||   |||:::|||:::||||||   |||||||||
Db        322 TACTTCAAGTTTATTCTCTGCAGATACACTTCAGTTTTGTTTTATGCAAACATGTATACT 381

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              ||||||:::   |||      |||||||||:::|||||||||   :::|||   ||||||
Db        382 TCCATCGTGTTCCTTGGGCTGATAAGCATTGATCGCTATCTGAAGGTGGTCAAGCCATTT 441

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
                  :::         |||:::      |||                  ||||||
Db        442 GGGGACTCTCGGATGTACAGCATAACCTTCACGAAGGTTTTATCTGTTTGTGTTTGGGTG 501

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
                 :::   :::      ||||||      ||||||                     |||
Db        502 ATCATGGCTGTTTTGTCTTTGCCAAACATCATCCTGACAAATGGT------------CAG 549

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
              |||      |||      ::::::|||:::      :::      ||||||
Db        550 CCA---------ACAGAGGACAATATCCATGACTGCTCAAAACTTAAAAGTCCTTTGGGG 600

Qy        279 LeuTyrTyrPheIleSerLeuAlaPhePheGly---------PheLeuIleProPheVal 295
              :::   :::      ::::::   :::              |||:::         |||
Db        601 GTCAAATGGCATACGGCAGTCACCTATGTGAACAGCTGCTTGTTTGTGGCCGTGCTGGTG 660

Qy        296 LeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrp 315
              :::::::|||   ||||||   ||||||   |||   ::::::              ::::::
Db        661 ATTCTGATCGGATGTTACATAGCCATATCCAGGTACATCCACAAATCCAGCAGGCAATTC 720

Qy        316 Leu---------------------TrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
              :::                     :::   ::::::      :::::::::   :::
Db        721 ATAAGTCAGTCAAGCCGAAAGCGAAAACATAACCAGAGCATCAGGGTTGTTGTGGCTGTG 780

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
              |||      ||||||   |||   ::::::      |||                 :::
Db        781 TTTTTTACCTGCTTTCTACCATATCACTTGTGCAGAATT------------CCTTTTACT 828

Qy        349 TyrAsnAsnThrAspGly--------------------------LeuTyrPheIleTyr 359
              :::::::::   |||                            ||||||:::
Db        829 TTTAGTCACTTAGACAGGCTTTTAGATGAATCTGCACAAAAAATCCTATATTACTGCAAA 888

Qy        360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379
                 |||   |||   |||   :::   |||   ||||||||||||   :::||||||
Db        889 GAAATTACACTTTTCTTGTCTGCGTGTAATGTTTGCCTGGATCCAATAATTTACTTTTTC 948

Qy        380 MetSerLysThr 383
              |||   ::::::
Db        949 ATGTGTAGGTCA 960

RESULT 40
US-10-225-567A-417
; Sequence 417, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 417
;   LENGTH: 1402
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-417

Alignment Scores:
Pred. No.:              8.98e-25      Length:          1402
Score:                  368.50        Matches:         89
Percent Similarity:     50.33%        Conservative:    64
Best Local Similarity:  29.28%        Mismatches:      116
Query Match:            17.08%        Indels:          35
DB:                     9             Gaps:            9

US-09-208-629F-6 (1-408) x US-10-225-567A-417 (1-1402)

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
              ||||||   |||         |||   :::         |||   :::::::|||   :::
Db        343 AAGAACACCACC---------CTTCACAATGAATTTGACACAATTGTCTTGCCGGTGCTT 393

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              ||||||::::::||||||      :::      |||   :::   :::|||   :::|||
Db        394 TATCTCATTATATTTGTGGCAAGCATCTTGCTGAATGGTTTAGCAGTGTGG---ATCTTC 450

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||      |||:::      :::   :::||||||      |||:::   :::||||||
Db        451 TTCCACATTAGGAATAAAACCAGCTTCATATTCTATCTCAAAAACATAGTGGTTGCAGAC 510

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
              :::      :::|||   ||||||:::|||  :::                    |||
Db        511 CTCATAATGACGCTGACATTTCCATTTCGAATAGTCCATGATGCAGGATTTGGACCTTGG 570

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              |||      :::|||||||||   |||:::|||:::||||||   |||||||||
Db        571 TACTTCAAGTTTATTCTCTGCAGATACACTTCAGTTTTGTTTTATGCAAACATGTATACT 630

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              ||||||:::   |||      |||||||||:::|||||||||   :::|||   ||||||
Db        631 TCCATCGTGTTCCTTGGGCTGATAAGCATTGATCGCTATCTGAAGGTGGTCAAGCCATTT 690
```

```
Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db        691 GGGGACTCTCGGATGTACAGCATAACCTTCACGAAGGTTTTATCTGTTTGTGTTTGGGTG 750

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db        751 ATCATGGCTGTTTTGTCTTTGCCAAACATCATCCTGACAAATGGT-----------CAG 798

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db        799 CCA---------ACAGAGGACAATATCCATGACTGCTCAAAACTTAAAAGTCCTTTGGGG 849

Qy        279 LeuTyrTyrPheIleSerLeuAlaPhePheGly---------PheLeuIleProPheVal 295
Db        850 GTCAAATGGCATACGGCAGTCACCTATGTGAACAGCTGCTTGTTTGTGGCCGTGCTGGTG 909

Qy        296 LeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrp 315
Db        910 ATTCTGATCGGATGTTACATAGCCATATCCAGGTACATCCACAAATCCAGCAGGCAATTC 969

Qy        316 Leu---------------------TrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db        970 ATAAGTCAGTCAAGCCGAAAGCGAAAACATAACCAGAGCATCAGGGTTGTTGTGGCTGTG 1029

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db       1030 TTTTTTACCTGCTTTCTACCATATCACTTGTGCAGAATT-----------CCTTTTACT 1077

Qy        349 TyrAsnAsnThrAspGly-------------------------LeuTyrPheIleTyr 359
Db       1078 TTTAGTCACTTAGACAGGCTTTTAGATGAATCTGCACAAAAAATCCTATATTACTGCAAA 1137

Qy        360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379
Db       1138 GAAATTACACTTTTCTTGTCTGCGTGTAATGTTTGCCTGGATCCAATAATTTACTTTTTC 1197

Qy        380 MetSerLysThr 383
Db       1198 ATGTGTAGGTCA 1209
```

Search completed: June 29, 2003, 11:46:58
Job time : 174.162 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:52:59 ; Search time 46.0564 Seconds
(without alignments)
2716.756 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIKM.....AYLTKXXNDLREQGQPSQRT 408

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09208629/runat_24062003_101626_12944/app_query.fasta_1.1166
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629_@CGN_1_1_47_@runat_24062003_101626_12944 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2142	99.3	1224	2	US-08-742-440A-4	Sequence 4, Appli
2	1947	90.3	1102	2	US-08-742-440A-5	Sequence 5, Appli
3	1443.5	66.9	1224	2	US-08-742-440A-1	Sequence 1, Appli
4	959	44.5	1124	2	US-08-742-440A-2	Sequence 2, Appli
5	606.5	28.1	2732	1	US-08-476-000-60	Sequence 60, Appl
6	606.5	28.1	2732	1	US-08-472-840-60	Sequence 60, Appl
7	606.5	28.1	2732	2	US-08-476-976-60	Sequence 60, Appl
8	606.5	28.1	2732	3	US-08-474-410-60	Sequence 60, Appl
9	606.5	28.1	2732	4	US-08-486-673B-60	Sequence 60, Appl
10	605.5	28.1	1475	1	US-08-097-938-1	Sequence 1, Appli
11	605.5	28.1	1475	1	US-08-476-000-1	Sequence 1, Appli
12	605.5	28.1	1475	1	US-08-472-840-1	Sequence 1, Appli
13	605.5	28.1	1475	2	US-08-476-976-1	Sequence 1, Appli
14	605.5	28.1	1475	3	US-08-474-410-1	Sequence 1, Appli
15	605.5	28.1	1475	4	US-08-486-673B-1	Sequence 1, Appli
16	585.5	27.1	1255	1	US-08-097-938-3	Sequence 3, Appli
17	585.5	27.1	1255	1	US-08-476-000-3	Sequence 3, Appli
18	585.5	27.1	1255	1	US-08-472-840-3	Sequence 3, Appli
19	585.5	27.1	1255	2	US-08-476-976-3	Sequence 3, Appli
20	585.5	27.1	1255	3	US-08-474-410-3	Sequence 3, Appli
21	585.5	27.1	1255	4	US-08-486-673B-3	Sequence 3, Appli
22	572.5	26.5	1414	1	US-08-476-000-62	Sequence 62, Appl
23	572.5	26.5	1414	1	US-08-472-840-62	Sequence 62, Appl
24	572.5	26.5	1414	2	US-08-476-976-62	Sequence 62, Appl
25	572.5	26.5	1414	3	US-08-474-410-62	Sequence 62, Appl
26	572.5	26.5	1414	4	US-08-486-673B-62	Sequence 62, Appl
27	523.5	24.3	4895	3	US-09-053-866-1	Sequence 1, Appli
28	523.5	24.3	4895	4	US-09-479-130-1	Sequence 1, Appli
29	523	24.2	3480	1	US-07-657-769B-68	Sequence 68, Appl
30	523	24.2	3480	1	US-07-789-184-219	Sequence 219, App
31	523	24.2	3480	1	US-08-475-263-219	Sequence 219, App
32	523	24.2	3480	1	US-08-485-886-219	Sequence 219, App
33	523	24.2	3480	2	US-08-477-362-219	Sequence 219, App
34	523	24.2	3480	2	US-08-477-134-219	Sequence 219, App
35	523	24.2	3480	3	US-08-473-489A-219	Sequence 219, App
36	523	24.2	3480	3	US-08-485-695-219	Sequence 219, App
37	523	24.2	3480	4	US-08-018-760-219	Sequence 219, App
38	516.5	23.9	1764	1	US-08-313-553-12	Sequence 12, Appl
39	516.5	23.9	1764	3	US-08-767-993-12	Sequence 12, Appl
40	410.5	19.0	1643	1	US-08-383-750-3	Sequence 3, Appli
41	410.5	19.0	1643	3	US-08-352-678-3	Sequence 3, Appli
42	410.5	19.0	1643	5	PCT-US93-09636-3	Sequence 3, Appli
43	410.5	19.0	1901	1	US-08-153-848-43	Sequence 43, Appl
44	410.5	19.0	1901	3	US-09-299-843A-43	Sequence 43, Appl
45	410.5	19.0	1901	4	US-09-088-337B-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-08-742-440A-4
; Sequence 4, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400

```
;      TELEFAX:  650 327-3231
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1224 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-742-440A-4

Alignment Scores:
Pred. No.:              3.14e-208      Length:         1224
Score:                  2142.00        Matches:        405
Percent Similarity:     99.26%         Conservative:   0
Best Local Similarity:  99.26%         Mismatches:     3
Query Match:            99.30%         Indels:         0
DB:                     2              Gaps:           0

US-09-208-629F-6 (1-408) x US-08-742-440A-4 (1-1224)

Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db          1 TGCTCCATGATTTTACAGATTTCATAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 60

Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 120

Qy         41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 180

Qy         61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 240

Qy         81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 300

Qy        101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 360

Qy        121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTCTTC 420

Qy        141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTTCTT 480

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTATTT 540

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 600

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTAC 660

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACAGTT 720

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 780

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTCTAT 840

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTACTGC 900

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 960

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATTCTT 1020

Qy        341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATATCTC 1080

Qy        361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTCATG 1140

Qy        381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGAGAA 1200

Qy        401 GlnGlyGlnProSerGlnArgThr 408
              |||||||||||||||||||||||||
Db       1201 CAAGGACAGCCATCACAGAGAACG 1224

RESULT 2
US-08-742-440A-5
; Sequence 5, Application US/08742440A
; Patent No. 5892014
;  GENERAL INFORMATION:
;    APPLICANT:  Coughlin, Shaun
;    APPLICANT:  Ishihari, Hiroaki
;    APPLICANT:  Connolly, Andrew
;    TITLE OF INVENTION:  Protease Activated Receptor
;    TITLE OF INVENTION:  3 and Uses Thereof
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bozicevic & Reed, LLP
;      STREET:  285 Hamilton Avenue, Suite 200
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,440A
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  36,677
;      REFERENCE/DOCKET NUMBER:  UCAL/060PAT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-327-3400
;      TELEFAX:  650 327-3231
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1102 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-742-440A-5

Alignment Scores:
Pred. No.:              1.66e-188       Length:         1102
Score:                  1947.00         Matches:        364
Percent Similarity:     99.46%          Conservative:   1
Best Local Similarity:  99.18%          Mismatches:     2
Query Match:            90.26%          Indels:         0
DB:                     2               Gaps:           0

US-09-208-629F-6 (1-408) x US-08-742-440A-5 (1-1102)

Qy         40 SerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACAGGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTT 60

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThr 79
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACA 120

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTG 180

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAAATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATC 240

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACCTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTC 300

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCAGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTT 360

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTTTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTA 420

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTTGGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCC 480

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTCTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACC 540

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACA 600

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTTTTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCA 660

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTC 720

Qy        280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TATTACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTAC 780

Qy        300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrVal 319
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGCTATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTT 840

Qy        320 LysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIle 339
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGGCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATT 900

Qy        340 LeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyr 359
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATAT 960

Qy        360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTCATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTC 1020

Qy        380 MetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArg 399
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATGTCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGA 1080

Qy        400 GluGlnGlyGlnProSerGln 406
              |||||||||||||||||||||
Db       1081 GAACAAGGACAGCCATCACAG 1101

RESULT 3
US-08-742-440A-1
; Sequence 1, Application US/08742440A
; Patent No. 5892014
;  GENERAL INFORMATION:
;    APPLICANT:  Coughlin, Shaun
;    APPLICANT:  Ishihari, Hiroaki
;    APPLICANT:  Connolly, Andrew
;    TITLE OF INVENTION:  Protease Activated Receptor
;    TITLE OF INVENTION:  3 and Uses Thereof
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bozicevic & Reed, LLP
;      STREET:  285 Hamilton Avenue, Suite 200
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,440A
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  36,677
;      REFERENCE/DOCKET NUMBER:  UCAL/060PAT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-327-3400
;      TELEFAX:  650 327-3231
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1224 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-742-440A-1

Alignment Scores:
Pred. No.:              2.45e-137       Length:         1224
Score:                  1443.50         Matches:        264
Percent Similarity:     83.74%          Conservative:   45
Best Local Similarity:  71.54%          Mismatches:     59
Query Match:            66.92%          Indels:         1
DB:                     2               Gaps:           1
```

US-09-208-629F-6 (1-408) x US-08-742-440A-1 (1-1224)

```
Qy          14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
Db          33 GGGTCTCAGGACATCAAGATGAAAATCCTTATCTTGGTTGCAGCTGGGCTGCTGTTTCTG 92

Qy          34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
Db          93 CCAGTCACTGTTTGCCAAAGTGGCATA---AATGTTTCAGACAACTCAGCAAAGCCAACC 149

Qy          54 LeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSer 73
Db         150 TTAACTATTAAGAGTTTTAATGGGGGTCCCCAAAATACCTTTGAAGAATTCCCACTTTCT 209

Qy          74 AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93
Db         210 GACATAGAGGGCTGGACAGGAGCCACCACAACTATAAAAGCGGAGTGTCCCGAGGACAGT 269

Qy          94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThr 113
Db         270 ATTTCAACTCTCCACGTGAATAATGCTACCATAGGATACCTGAGAAGTTCCTTAAGTACC 329

Qy         114 LysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaVal 133
Db         330 CAAGTGATACCTGCCATCTATATCCTGCTGTTTGTGGTTGGTGTACCATCCAACATCGTG 389

Qy         134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153
Db         390 ACCCTGTGGAAACTCTCCTTAAGGACCAAATCCATCAGTCTGGTCATCTTTCACACCAAC 449

Qy         154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
Db         450 CTGGCCATCGCAGATCTCCTTTTCTGTGTCACACTGCCATTTAAGATCGCCTACCATCTC 509

Qy         174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
Db         510 AATGGCAACAACTGGGTATTTGGCGAGGTCATGTGCCGGATCACCACGGTCGTTTTCTAC 569

Qy         194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
Db         570 GGCAACATGTACTGCGCTATCCTGATCCTCACTTGCATGGGCATCAACCGCTACCTGGCC 629

Qy         214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233
Db         630 ACGGCTCACCCTTTCACATACCAGAAGCTGCCCAAACGCAGCTTCTCCTTGCTCATGTGT 689

Qy         234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
Db         690 GGCATAGTGTGGGTCATGGTTTTCTTATACATGCTGCCCTTTGTCATCCTGAAGCAGGAG 749

Qy         254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
Db         750 TACCACCTCGTCCACTCAGAGATCACCACCTGCCACGATGTCGTCGACGCGTGCGAGTCC 809

Qy         274 SerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIlePro 293
Db         810 CCATCATCCTTCCGATTCTACTACTTCGTCTCCTTAGCATTCTTTGGGTTCCTCATCCCG 869

Qy         294 PheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHis 313
Db         870 TTTGTGATCATCATCTTCTGTTACACGACTCTCATCCACAAACTTAAATCAAAGGATCGG 929

Qy         314 ArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPhe 333
Db         930 ATATGGCTGGGCTACATCAAGGCCGTCCTCCTCATCCTTGTGATTTTCACAATTTGCTTT 989

Qy         334 AlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAsp 353
Db         990 GCCCCCACCAACATCATACTCGTAATCCACCATGCCAACTACTACTACCACAATACCGAC 1049

Qy         354 GlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
```

```
Db        1050 AGCTTGTACTTTATGTATCTTATTGCTCTGTGCCTGGGGAGCCTGAATAGCTGCCTAGAT 1109

Qy         374 ProPheLeuTyrPheLeuMetSerLys 382
Db        1110 CCATTCCTTTACTTTGTCATGTCGAAA 1136
```

RESULT 4
US-08-742-440A-2
; Sequence 2, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
US-08-742-440A-2

Alignment Scores:

Pred. No.:	3.22e-88	Length:	1124
Score:	959.00	Matches:	185
Percent Similarity:	73.99%	Conservative:	34
Best Local Similarity:	62.50%	Mismatches:	53
Query Match:	44.46%	Indels:	25
DB:	2	Gaps:	2

US-09-208-629F-6 (1-408) x US-08-742-440A-2 (1-1124)

```
Qy          44 AsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaPro 63
Db          36 AATGTTTCAGACAACTCAGCAAAGCCAACCTTAACTATTAAGAGTTTTAATGGGGGTCCC 95

Qy          64 ProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIle 83
Db          96 CAAAATACCTTTGAAGAATTCNN---------------------------------- 118

Qy          84 ThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThr 103
```

```
Db        119 --------------------------NNNNNTACAACTCTCCATGTGAATAATGCTACC 151

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123

Db        152 ATGGGATACCTGAGAAGTTCCTTAAGTACCAAAGTGATACCTGCCATCTACATCCTGGTG 211

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArg 143

Db        212 TTTGTGATTGGTGTACCAGCGAACATCGTGACCCTGTGGAAACTCTCCTCAAGGACCAAA 271

Qy        144 SerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163

Db        272 TCCATCTGTCTGGTCATCTTTCACACCAACCTGGCCATCGCGGATCTCCTTTTCTGTGTC 331

Qy        164 ThrLeuProPheLysIle-AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVa 183

Db        332 ACGCTGCCGTTTAAGATCNNCCTACCATCTCAATGGCAACAACTGGGTATTTGGCGAGGT 391

Qy        183 lLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer---IleLeuLe 202

Db        392 CATGTGCCGGATCACCACGGTCGTTTTCTACGGCAACATGTACTGCGCTANNNTCCTGAT 451

Qy        202 uLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGl 222

Db        452 CCTCACCTGCATGGGCATCAACCGCTACCTGGCCACGGCTCACCCTTTCACATACCAGAA 511

Qy        222 yLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLe 242

Db        512 GCTGCCCAAACGCAGCTTCTCCATGCTCATGTGTGGCATGGTGTGGGTCATGGTTTTCTT 571

Qy        242 uTyrMetLeuProPhePheIle-LeuLysGlnGluTyrTyrLeuValGlnProAspIleT 262

Db        572 ATACATGCTGCCCTTTGTCATCCNNNAAGCAGGAGTACCACCTCGTCCACTCCGAGATCA 631

Qy        262 hrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrP 282

Db        632 CCACCTGCCACGATGTCGTCGACGCGTGCGANTCCCCATCATCCTTCCGATTCTACTACT 691

Qy        282 heIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrA 302

Db        692 TCGTCTCCTTAGCATTCTTTGGGTTCCTCATCCCGTTTGTGATCATCATCTTCTGTTACA 751

Qy        302 laAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaS 322

Db        752 CGACTCTCATCCACAAACTTAAATCAAAAGATCNGATATGGCTGGGCTACATCAAGGCCG 811

Qy        322 erLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336

Db        812 TCCTCCTCATCCTTGTGAATTTCACCATCTGCTTCCCCCCCACC 855

RESULT 5
US-08-476-000-60
; Sequence 60, Application US/08476000
; Patent No. 5716789
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,000
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2732 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  73..1269
US-08-476-000-60

Alignment Scores:
Pred. No.:                6.99e-52        Length:          2732
Score:                    606.50          Matches:         137
Percent Similarity:       53.23%          Conservative:    69
Best Local Similarity:    35.40%          Mismatches:      126
Query Match:              28.12%          Indels:          55
DB:                       1               Gaps:            12

US-09-208-629F-6 (1-408) x US-08-476-000-60 (1-2732)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41

Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59

Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThr 79

Db        205 GAAACCCAGCCTCCA-----------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97

Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117

Db        280 ----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137

Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156

Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176

Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196

Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
```

```
Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235

Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253

Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273

Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG---------------- 765

Qy        274 SerSerProPheGlnLeuTyr----------------TyrPheIleSerLeuAlaPhe 287

Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307

Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321

Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341

Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361

Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381

Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387

Db       1123 AAAGATTTCAGGGATCACGCC 1143

RESULT 6
US-08-472-840-60
; Sequence 60, Application US/08472840
; Patent No. 5763575
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,840
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2732 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  73..1269
US-08-472-840-60

Alignment Scores:
Pred. No.:                6.99e-52        Length:          2732
Score:                    606.50          Matches:         137
Percent Similarity:       53.23%          Conservative:    69
Best Local Similarity:    35.40%          Mismatches:      126
Query Match:              28.12%          Indels:          55
DB:                       1               Gaps:            12

US-09-208-629F-6 (1-408) x US-08-472-840-60 (1-2732)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41

Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59

Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThr 79

Db        205 GAAACCCAGCCTCCA----------------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97

Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117

Db        280 ----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137

Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156

Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176

Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196

Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216

Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235

Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
```

```
Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273

Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG---------------------- 765

Qy        274 SerSerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhe 287

Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307

Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla-----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321

Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341

Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361

Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381

Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387

Db       1123 AAAGATTTCAGGGATCACGCC 1143

RESULT 7
US-08-476-976-60
; Sequence 60, Application US/08476976
; Patent No. 5874400
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,976
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2732 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  73..1269
US-08-476-976-60

Alignment Scores:
Pred. No.:              6.99e-52        Length:         2732
Score:                  606.50          Matches:        137
Percent Similarity:     53.23%          Conservative:   69
Best Local Similarity:  35.40%          Mismatches:     126
Query Match:            28.12%          Indels:         55
DB:                     2               Gaps:           12

US-09-208-629F-6 (1-408) x US-08-476-976-60 (1-2732)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41

Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59

Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThr 79

Db        205 GAAACCCAGCCTCCA-------------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97

Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117

Db        280 ----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137

Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156

Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176

Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196

Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216

Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235

Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253

Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273

Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG---------------------- 765

Qy        274 SerSerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhe 287
```

```
Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla-----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321
Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387
Db       1123 AAAGATTTCAGGGATCACGCC 1143

RESULT 8
US-08-474-410-60
; Sequence 60, Application US/08474410
; Patent No. 6043212
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,410
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2732 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  73..1269
```

US-08-474-410-60

Alignment Scores:

Pred. No.:	6.99e-52	Length:	2732
Score:	606.50	Matches:	137
Percent Similarity:	53.23%	Conservative:	69
Best Local Similarity:	35.40%	Mismatches:	126
Query Match:	28.12%	Indels:	55
DB:	3	Gaps:	12

US-09-208-629F-6 (1-408) x US-08-474-410-60 (1-2732)

```
Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
Db        205 GAAACCCAGCCTCCA-----------------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97
Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
Db        280 -----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137
Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG----------------- 765

Qy        274 SerSerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhe 287
Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla-----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321
```

```
Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387
Db       1123 AAAGATTTCAGGGATCACGCC 1143

RESULT 9
US-08-486-673B-60
; Sequence 60, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
;  APPLICANT: Sundelin, Johan
;  APPLICANT:  Scarborough, Robert M.
;  TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
;  FILE REFERENCE: 44481-5006-08-US
;  CURRENT APPLICATION NUMBER: US/08/486,673B
;  CURRENT FILING DATE:  1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/097,938
;  PRIOR FILING DATE: 1993-07-26
;  PRIOR APPLICATION NUMBER: PCT/US94/08536
;  PRIOR FILING DATE: 1994-07-26
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
;   LENGTH: 2732
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (73)..(1269)
;   OTHER INFORMATION: C140 receptor, cDNA and deduced protein sequences
US-08-486-673B-60

Alignment Scores:
Pred. No.:              6.99e-52      Length:          2732
Score:                  606.50        Matches:         137
Percent Similarity:     53.23%        Conservative:    69
Best Local Similarity:  35.40%        Mismatches:      126
Query Match:            28.12%        Indels:          55
DB:                     4             Gaps:            12

US-09-208-629F-6 (1-408) x US-08-486-673B-60 (1-2732)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThr 79
Db        205 GAAACCCAGCCTCCA-----------------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97
Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
Db        280 -----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137
Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG---------------------- 765

Qy        274 SerSerProPheGlnLeuTyr------------------TyrPheIleSerLeuAlaPhe 287
Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla------------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321
Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387
Db       1123 AAAGATTTCAGGGATCACGCC 1143

RESULT 10
US-08-097-938-1
; Sequence 1, Application US/08097938
; Patent No. 5629174
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR AND ITS
;    TITLE OF INVENTION:  AGONISTS AND ANTAGONISTS
;    NUMBER OF SEQUENCES:  59
```

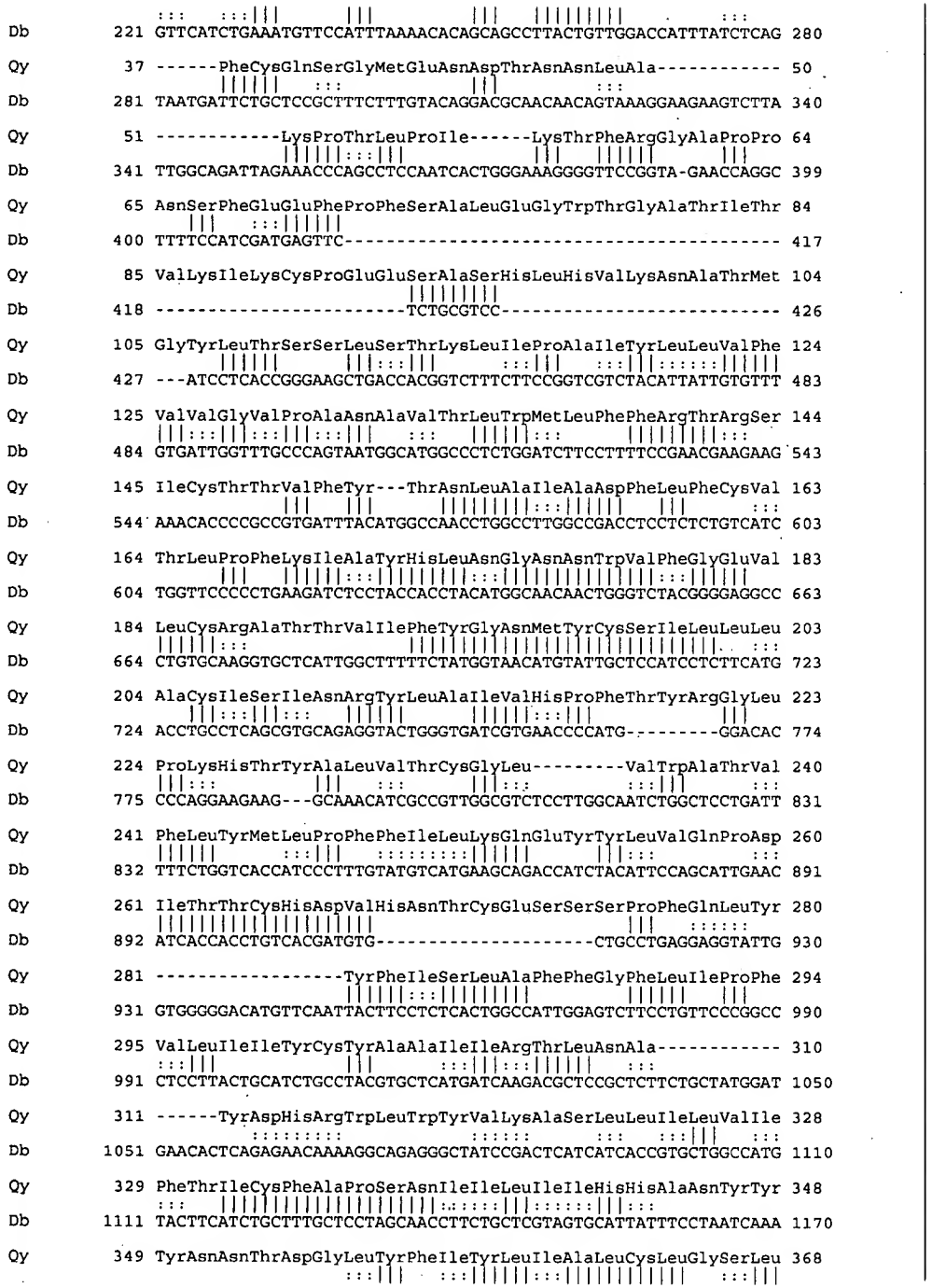
```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/097,938
;      FILING DATE:  26-JUL-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22803-20006.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1475 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  232..1416
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  232
US-08-097-938-1

Alignment Scores:
Pred. No.:                3.61e-52        Length:          1475
Score:                    605.50          Matches:         140
Percent Similarity:       52.50%          Conservative:    70
Best Local Similarity:    35.00%          Mismatches:      117
Query Match:              28.07%          Indels:          74
DB:                       1               Gaps:            13

US-09-208-629F-6 (1-408) x US-08-097-938-1 (1-1475)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla---------- 50
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 ----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db        400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db        418 ----------------------TCTGCGTCC-------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        892 ATCACCACCTGTCACGATGTG----------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla---------- 310
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db       1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290

RESULT 11
US-08-476-000-1
; Sequence 1, Application US/08476000
; Patent No. 5716789
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

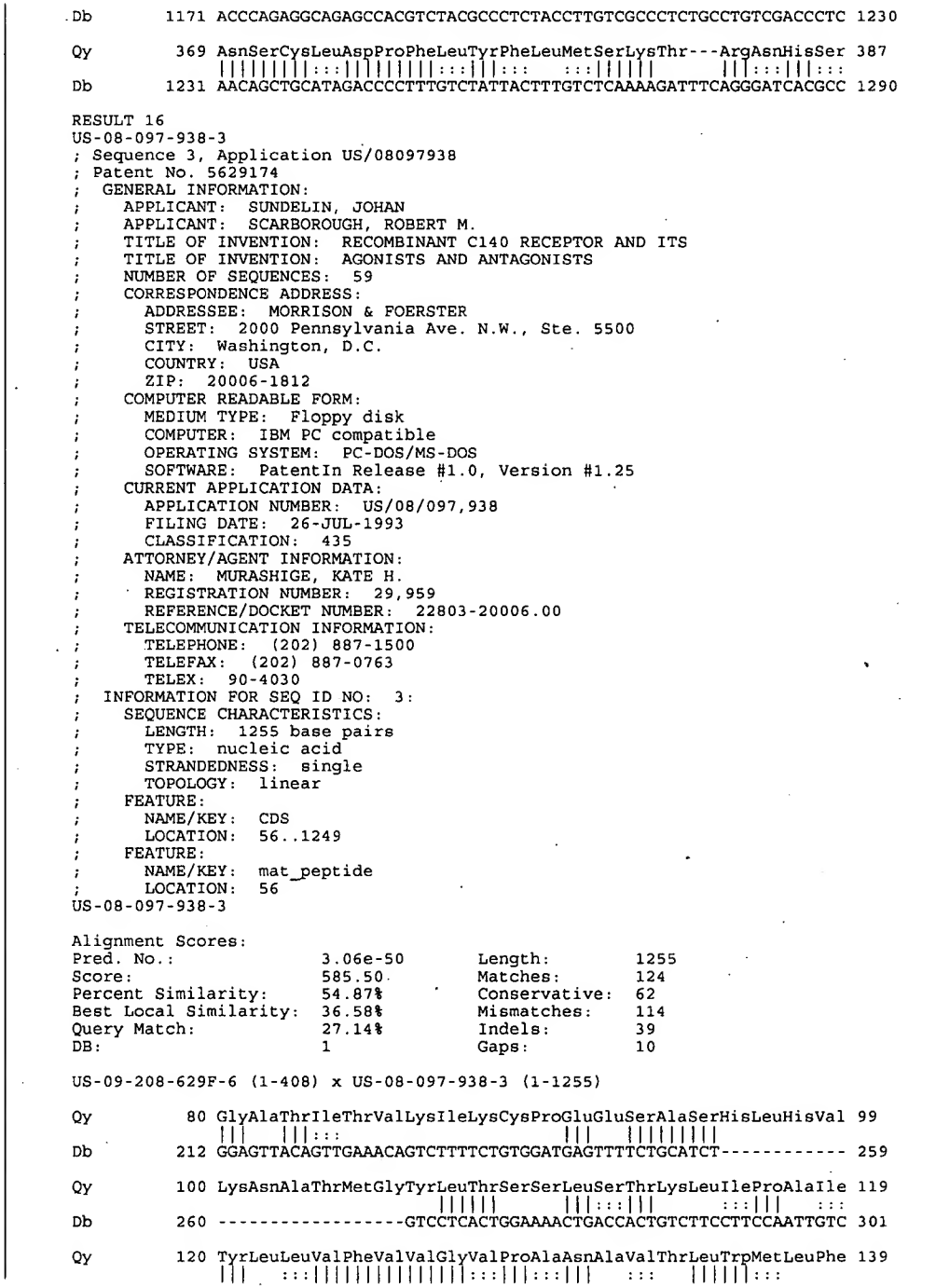
```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/476,000
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/390,301
;       FILING DATE:  25-JAN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  ADLER, REID G.
;       REGISTRATION NUMBER:  30,988
;       REFERENCE/DOCKET NUMBER:  2803-0006.20
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 887-1500
;       TELEFAX:  (202) 887-0763
;       TELEX:  90-4030
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1475 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  232..1416
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  232
US-08-476-000-1

Alignment Scores:
Pred. No.:                3.61e-52        Length:          1475
Score:                    605.50          Matches:         140
Percent Similarity:       52.50%          Conservative:    70
Best Local Similarity:    35.00%          Mismatches:      117
Query Match:              28.07%          Indels:          74
DB:                       1               Gaps:            13

US-09-208-629F-6 (1-408) x US-08-476-000-1 (1-1475)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla---------- 50
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 ----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db        400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db        418 ----------------------TCTGCGTCC-------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        892 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db       1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290

RESULT 12
US-08-472-840-1
; Sequence 1, Application US/08472840
; Patent No. 5763575
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,840
```

```
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/390,301
;       FILING DATE:  25-JAN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  ADLER, REID G.
;       REGISTRATION NUMBER:  30,988
;       REFERENCE/DOCKET NUMBER:  2803-0006.20
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 887-1500
;       TELEFAX:  (202) 887-0763
;       TELEX:  90-4030
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1475 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  232..1416
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  232
US-08-472-840-1

Alignment Scores:
Pred. No.:                3.61e-52        Length:          1475
Score:                    605.50          Matches:         140
Percent Similarity:       52.50%          Conservative:    70
Best Local Similarity:    35.00%          Mismatches:      117
Query Match:              28.07%          Indels:          74
DB:                       1               Gaps:            13

US-09-208-629F-6 (1-408) x US-08-472-840-1 (1-1475)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuLeuProThr--- 36
              :::  :::|||     |||       |||  |||||||||          :::
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla------------ 50
                  ||||||   :::       |||         :::
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 ------------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
                       ||||||:::|||            |||   ||||||      |||
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
              |||   :::||||||
Db        400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
                                    |||||||||
Db        418 ----------------------TCTGCGTCC---------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
                  ||||||      |||:::|||      :::|||   :::|||::::::||||||
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
              |||:::|||:::|||:::|||   :::   ||||||:::      |||||||||:::
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
                     |||   |||      |||||||||:::||||||   |||      :::
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
                  |||   ||||||:::|||||||||:::|||||||||||||||:::|||||||
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
              ||||||:::                |||||||||||||||||||||||||||||   :::
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
                 |||:::|||:::   ||||||      ||||||:::|||            |||
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
              |||:::         |||   :::      |||:::         :::|||      :::
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||      :::|||   :::::::::||||||      |||:::          :::
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||                              |||   ::::::
Db        892 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
                              ||||||:::|||||||||         ||||||   |||
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
              :::|||            |||      :::|||:::||||||   :::
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
                    :::::::::          ::::::      :::   :::|||   :::
Db       1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
              :::   ||||||||||||||||||||:::::|||::::::|||:::
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
                                  :::|||   :::|||||||:::||||||||||   :::|||
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
              |||||||||:::|||||||||:::|||:::   :::||||||      |||:::|||:::
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290

RESULT 13
US-08-476-976-1
; Sequence 1, Application US/08476976
; Patent No. 5874400
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,976
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1475 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  232..1416
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  232
US-08-476-976-1

Alignment Scores:
Pred. No.:                3.61e-52        Length:          1475
Score:                    605.50          Matches:         140
Percent Similarity:       52.50%          Conservative:    70
Best Local Similarity:    35.00%          Mismatches:      117
Query Match:              28.07%          Indels:          74
DB:                       2               Gaps:            13

US-09-208-629F-6 (1-408) x US-08-476-976-1 (1-1475)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla---------- 50
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 ----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db        400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db        418 ----------------------TCTGCGTCC-------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        892 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db       1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290

RESULT 14
US-08-474-410-1
; Sequence 1, Application US/08474410
; Patent No. 6043212
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,410
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1475 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  232..1416
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  232
US-08-474-410-1

Alignment Scores:
Pred. No.:              3.61e-52        Length:         1475
Score:                  605.50          Matches:        140
Percent Similarity:     52.50%          Conservative:   70
Best Local Similarity:  35.00%          Mismatches:     117
Query Match:            28.07%          Indels:         74
DB:                     3               Gaps:           13

US-09-208-629F-6 (1-408) x US-08-474-410-1 (1-1475)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla---------- 50
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 ----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db        400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db        418 ------------------------TCTGCGTCC-------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        892 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPheGlyPheLeuIleProPhe 294
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla---------- 310
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db       1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290

RESULT 15
US-08-486-673B-1
; Sequence 1, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
;  APPLICANT: Sundelin, Johan
;  APPLICANT:  Scarborough, Robert M.
;  TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
;  FILE REFERENCE: 44481-5006-08-US
;  CURRENT APPLICATION NUMBER: US/08/486,673B
;  CURRENT FILING DATE:  1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/097,938
;  PRIOR FILING DATE: 1993-07-26
;  PRIOR APPLICATION NUMBER: PCT/US94/08536
;  PRIOR FILING DATE: 1994-07-26
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (232)..(1416)
;   OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein
;   OTHER INFORMATION: sequences
US-08-486-673B-1

Alignment Scores:
Pred. No.:              3.61e-52        Length:         1475
Score:                  605.50          Matches:        140
Percent Similarity:     52.50%          Conservative:   70
Best Local Similarity:  35.00%          Mismatches:     117
Query Match:            28.07%          Indels:         74
DB:                     4               Gaps:           13

US-09-208-629F-6 (1-408) x US-08-486-673B-1 (1-1475)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
```





RESULT 16
US-08-097-938-3
; Sequence 3, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,938
; FILING DATE: 26-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22803-20006.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..1249
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 56
US-08-097-938-3

Alignment Scores:

Pred. No.:	3.06e-50	Length:	1255
Score:	585.50	Matches:	124
Percent Similarity:	54.87%	Conservative:	62
Best Local Similarity:	36.58%	Mismatches:	114
Query Match:	27.14%	Indels:	39
DB:	1	Gaps:	10

US-09-208-629F-6 (1-408) x US-08-097-938-3 (1-1255)

```
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                |||   |||      |||||||||:::||||||
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
                 |||      :::      |||   ||||||||||||||:::::||||||||||||
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              ::::::||||||   ||||||                  ||||||||||||||||||||
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
                 :::          |||          |||:::                 :::|||
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
                 :::   |||      :::|||   :::::::::||||||      ::::::
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
                 :::||||||||||||||||||||                          |||   |||
Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              |||                  |||||:::|||||||||          ||||||
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
              |||       |||            |||       :::||||||   |||   :::
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                        :::   :::            :::|||      :::   :::|||
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
              ::::::   |||||||||   |||||||||:::::|||:::::|||
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
              |||:::          :::   |||  .      :::|||   :::|||::::::||||||
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   :::|||||||||||||:::|||||||||:::|||:::   :::|||
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                 |||:::|||:::      |||   |||   :::           :::::::::
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 17
US-08-476-000-3
; Sequence 3, Application US/08476000
; Patent No. 5716789
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,000
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1255 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  56..1249
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  56
US-08-476-000-3

Alignment Scores:
Pred. No.:              3.06e-50        Length:         1255
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     1               Gaps:           10

US-09-208-629F-6 (1-408) x US-08-476-000-3 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||   |||:::                      |||   |||||||||
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                                  ||||||      |||:::|||      :::|||   :::
Db        260 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   :::|||||||||||||||:::|||:::|||   :::   ||||||:::
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                |||   |||      |||||||||:::||||||
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
                 |||      :::      |||   ||||||||||||||:::::||||||||||||
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              ::::::||||||   ||||||                  ||||||||||||||||||||
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601
```

```
Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 ------------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165
```

```
RESULT 18
US-08-472-840-3
; Sequence 3, Application US/08472840
; Patent No. 5763575
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,840
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1255 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  56..1249
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  56
US-08-472-840-3

Alignment Scores:
Pred. No.:                3.06e-50      Length:          1255
Score:                    585.50        Matches:         124
Percent Similarity:       54.87%        Conservative:    62
Best Local Similarity:    36.58%        Mismatches:      114
Query Match:              27.14%        Indels:          39
DB:                       1             Gaps:            10

US-09-208-629F-6 (1-408) x US-08-472-840-3 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
Db        260 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817
```

```
Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 19
US-08-476-976-3
; Sequence 3, Application US/08476976
; Patent No. 5874400
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,976
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1255 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  56..1249
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  56
```

US-08-476-976-3

Alignment Scores:

Pred. No.:	3.06e-50	Length:	1255
Score:	585.50	Matches:	124
Percent Similarity:	54.87%	Conservative:	62
Best Local Similarity:	36.58%	Mismatches:	114
Query Match:	27.14%	Indels:	39
DB:	2	Gaps:	10

US-09-208-629F-6 (1-408) x US-08-476-976-3 (1-1255)

```
Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
Db        260 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
```

```
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 20
US-08-474-410-3
; Sequence 3, Application US/08474410
; Patent No. 6043212
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,410
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1255 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  56..1249
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  56
US-08-474-410-3

Alignment Scores:
Pred. No.:              3.06e-50      Length:          1255
Score:                  585.50        Matches:         124
Percent Similarity:     54.87%        Conservative:    62
Best Local Similarity:  36.58%        Mismatches:      114
Query Match:            27.14%        Indels:          39
DB:                     3             Gaps:            10

US-09-208-629F-6 (1-408) x US-08-474-410-3 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119

Db        260 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 ------------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 21
US-08-486-673B-3
; Sequence 3, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
;  APPLICANT: Sundelin, Johan
;  APPLICANT:  Scarborough, Robert M.
;  TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
;  FILE REFERENCE: 44481-5006-08-US
;  CURRENT APPLICATION NUMBER: US/08/486,673B
```

```
;  CURRENT FILING DATE:  1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/097,938
;  PRIOR FILING DATE: 1993-07-26
;  PRIOR APPLICATION NUMBER: PCT/US94/08536
;  PRIOR FILING DATE: 1994-07-26
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1255
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (56)..(1249)
;   OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein
;   OTHER INFORMATION: sequences
US-08-486-673B-3

Alignment Scores:
Pred. No.:              3.06e-50      Length:          1255
Score:                  585.50        Matches:         124
Percent Similarity:     54.87%        Conservative:    62
Best Local Similarity:  36.58%        Mismatches:      114
Query Match:            27.14%        Indels:          39
DB:                     4             Gaps:            10

US-09-208-629F-6 (1-408) x US-08-486-673B-3 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||   |||:::                  |||   |||||||||
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT---------- 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                                 ||||||      |||:::|||      :::|||   :::
Db        260 ------------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   :::|||||||||||||||:::|||:::|||   :::   ||||||:::
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                  |||   |||      |||||||||:::||||||
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
                 |||      :::      |||   ||||||||||||||||:::::|||||||||||
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              ::::::||||||   ||||||                  ||||||||||||||||||||
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
                 :::              |||         |||:::                :::|||
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
                 :::   |||      :::|||   :::::::::||||||      ::::::
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
                 :::||||||||||||||||||||                        |||   |||
Db        719 CTGAACATCACGACCTGTCATGATGTT----------*---------TTGCCTGAGCAG 757

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              |||                  ||||||:::|||||||||         ||||||
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
                 |||      |||              |||      :::||||||   |||   :::
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                          :::   :::            :::|||      :::   :::|||
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
                 ::::::   |||||||||   |||||||||:::::|||:::::|||
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
              |||:::          :::   |||         :::|||   :::|||::::::||||||
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   :::|||||||||||||:::||||||||||:::|||:::   :::|||
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                 |||:::|||:::      |||   |||   :::          :::::::::
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 22
US-08-476-000-62
; Sequence 62, Application US/08476000
; Patent No. 5716789
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,000
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  62:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1414 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  50..1240
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  50
US-08-476-000-62
```

```
Alignment Scores:
Pred. No.:                7.59e-49        Length:          1414
Score:                    572.50          Matches:         140
Percent Similarity:       49.64%          Conservative:    66
Best Local Similarity:    33.73%          Mismatches:      152
Query Match:              26.54%          Indels:          57
DB:                       1               Gaps:            14

US-09-208-629F-6 (1-408) x US-08-476-000-62 (1-1414)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
              |||   |||:::|||      :::           :::   |||   |||      |||
Db         41 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 100

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
              |||   |||                    |||   ||||||          ||||||
Db        101 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 139

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
                 :::|||          |||          |||
Db        140 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA-------------- 184

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
                          |||   :::  |||      |||   |||:::
Db        185 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                    |||   |||||||||                                |||
Db        227 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 256

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
                   |||:::|||      :::|||   :::|||   :::||||||   ||||||
Db        257 GCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGT 316

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
              :::|||:::|||   :::   |||||:::      ||||||||:::
Db        317 TTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTCCGAACTAAGAAGAAGCACCCT 376

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||  |||      |||||||||:::||||||   |||      :::      |||
Db        377 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 436

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
                 ||||||||||||||:::::|||||||||||:::::||||||   ||||||
Db        437 TTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 496

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                       ||||||   ||||||||||||||||||||   :::   |||:::
Db        497 GTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 556

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
              |||:::   ||||||      ||||||:::|||      :::              |||
Db        557 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 613

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
                 |||:::                 :::|||              |||      :::|||
Db        614 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCT 673

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
                 :::::::::||||||      ::::::         :::||||||||||||||||||
Db        674 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 733

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr----------------- 280
              |||                          |||   ||||||
Db        734 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||:::|||||||||          ||||||   |||      |||
Db        773 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 832

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla-----------------TyrAspHisArg 314
              |||      :::||||||  .|||   :::                      :::   :::

Db        833 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
                       :::|||      :::   :::|||   ::::::   |||||||||
Db        893 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACT 952

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
                 |||||||||::::::|||::::::|||          |||:::          :::   |||
Db        953 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1003

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
                        :::|||   :::|||::::::|||||||||||||   :::||||||||||||
Db       1004 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1063

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
                 :::||||||||:::|||:::   :::|||          |||:::|||:::      |||
Db       1064 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1123

Qy        390 TyrLeuThrLys******AsnAspLeuArgGluGlnGlyGlnPro 404
                 |||   :::           :::::::::          |||
Db       1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168

RESULT 23
US-08-472-840-62
; Sequence 62, Application US/08472840
; Patent No. 5763575
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,840
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  62:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1414 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  50..1240
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  50
US-08-472-840-62
```

```
Alignment Scores:
Pred. No.:                7.59e-49        Length:          1414
Score:                    572.50          Matches:         140
Percent Similarity:       49.64%          Conservative:    66
Best Local Similarity:    33.73%          Mismatches:      152
Query Match:              26.54%          Indels:          57
DB:                       1               Gaps:            14

US-09-208-629F-6 (1-408) x US-08-472-840-62 (1-1414)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
              |||   |||:::|||     :::           :::  |||   |||      |||
Db         41 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 100

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
              |||   |||                  |||   ||||||              ||||||
Db        101 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 139

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
                 :::|||           |||           |||
Db        140 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA--------------- 184

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
                          |||   :::   |||     |||  |||:::
Db        185 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                    |||   |||||||||                                   |||
Db        227 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 256

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
                    |||:::|||      :::|||   :::|||   :::||||||   ||||||
Db        257 GCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGT 316

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
              :::|||:::|||   :::   |||||:::      ||||||||:::
Db        317 TTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 376

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||   |||      |||||||||:::||||||   |||      :::      |||
Db        377 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 436

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
                 ||||||||||||||||:::::||||||||||:::::||||||   ||||||
Db        437 TTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 496

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                       ||||||   ||||||||||||||||||||   :::   |||:::
Db        497 GTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 556

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
              |||:::   ||||||     ||||||:::|||      :::          |||
Db        557 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 613

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
                 |||:::               :::|||            |||      :::|||
Db        614 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCT 673

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
                 :::::::::||||||     ::::::      :::||||||||||||||||||
Db        674 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 733

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr----------------- 280
              |||                        |||   ||||||
Db        734 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772

Qy        281 TyrPheIleSerLeuAlaPheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||:::|||||||||          ||||||   |||      |||
Db        773 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 832

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla-----------------TyrAspHisArg 314
              |||     :::||||||  |||   :::                  :::   :::
Db        833 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
                          :::|||      :::   :::|||   ::::::   |||||||||
Db        893 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACT 952

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
              |||||||||:::::::|||:::::::|||          |||:::        :::   |||
Db        953 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1003

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
                       :::|||   :::|||::::::|||||||||||   :::||||||||||||
Db       1004 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1063

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
              :::|||||||||:::|||:::   :::|||          |||:::|||:::      |||
Db       1064 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1123

Qy        390 TyrLeuThrLys******AsnAspLeuArgGluGlnGlyGlnPro 404
                 |||   :::           :::::::::        |||
Db       1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168

RESULT 24
US-08-476-976-62
; Sequence 62, Application US/08476976
; Patent No. 5874400
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,976
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  62:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1414 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  50..1240
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  50
US-08-476-976-62

Alignment Scores:
```

```
Pred. No.:              7.59e-49        Length:         1414
Score:                  572.50          Matches:        140
Percent Similarity:     49.64%          Conservative:   66
Best Local Similarity:  33.73%          Mismatches:     152
Query Match:            26.54%          Indels:         57
DB:                     2               Gaps:           14

US-09-208-629F-6 (1-408) x US-08-476-976-62 (1-1414)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
              |||   |||:::|||     :::         :::   |||   |||     |||
Db         41 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 100

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
              |||   |||                   |||   ||||||            ||||||
Db        101 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 139

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
                 :::|||           |||           |||
Db        140 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA-------------- 184

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
                       |||   :::   |||     |||   |||:::
Db        185 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                    |||   |||||||||                                   |||
Db        227 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 256

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
                  |||:::|||       :::|||   :::|||   :::||||||   ||||||
Db        257 GCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGT 316

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
              :::|||:::|||   :::   ||||||:::       ||||||||:::
Db        317 TTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 376

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||   |||       |||||||||:::||||||   |||       :::       |||
Db        377 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 436

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
                  ||||||||||||||:::::||||||||||:::::||||||   ||||||
Db        437 TTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 496

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                       ||||||   ||||||||||||||||||||   :::   |||:::
Db        497 GTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 556

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
              |||:::   ||||||      ||||||:::|||       :::             |||
Db        557 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 613

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
                 |||:::                 :::|||           |||       ::::|||
Db        614 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCT 673

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
                 ::::::::::||||||      ::::::         :::||||||||||||||||
Db        674 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 733

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr------------------ 280
              |||                         |||   ||||||
Db        734 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||:::|||||||||          ||||||   |||       |||
Db        773 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 832

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla----------------TyrAspHisArg 314
              |||      :::||||||   |||   :::                     :::   :::
Db        833 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
                          :::|||       :::   :::|||   ::::::   |||||||||
Db         893 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACT 952

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
              |||||||||:::::|||:::::|||          |||:::          :::   |||
Db        953 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1003

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
                      :::|||   :::|||::::::||||||||||||   :::||||||||||||
Db       1004 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1063

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
              :::|||||||||:::|||:::   :::|||          |||:::|||:::      |||
Db       1064 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1123

Qy        390 TyrLeuThrLys******AsnAspLeuArgGluGlnGlyGlnPro 404
                 |||   :::           :::::::::         |||
Db       1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168

RESULT 25
US-08-474-410-62
; Sequence 62, Application US/08474410
; Patent No. 6043212
;  GENERAL INFORMATION:
;    APPLICANT:  SUNDELIN, JOHAN
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT C140 RECEPTOR, ITS AGONISTS
;    TITLE OF INVENTION:  AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave. N.W., Ste. 5500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,410
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/390,301
;      FILING DATE:  25-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ADLER, REID G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0006.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  62:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1414 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  50..1240
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  50
US-08-474-410-62

Alignment Scores:
Pred. No.:              7.59e-49        Length:         1414
```

```
Score:                  572.50          Matches:        140
Percent Similarity:     49.64%          Conservative:   66
Best Local Similarity:  33.73%          Mismatches:     152
Query Match:            26.54%          Indels:         57
DB:                     3               Gaps:           14

US-09-208-629F-6 (1-408) x US-08-474-410-62 (1-1414)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
              |||   |||:::|||      :::           :::   |||   |||      |||
Db         41 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 100

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
              |||   |||                 |||   ||||||                ||||||
Db        101 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 139

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
                 :::|||           |||           |||
Db        140 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA-------------- 184

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
                          |||   :::   |||      |||   .|||:::
Db        185 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                    |||   |||||||||                                    |||
Db        227 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 256

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
                  |||:::|||      :::|||   ::::||   :::||||||   ||||||
Db        257 GCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGT 316

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
              :::|||:::|||   :::   ||||||:::      ||||||||||:::
Db        317 TTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 376

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||   |||      |||||||||:::||||||   |||      :::      |||
Db        377 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 436

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
                 ||||||||||||||:::::::||||||||||:::::::||||||   ||||||
Db        437 TTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 496

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                       ||||||   ||||||||||||||||||||||   :::   |||:::
Db        497 GTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 556

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
              |||:::   ||||||      ||||||:::|||      :::            |||
Db        557 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 613

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
                    |||:::                 :::|||            |||      :::|||
Db        614 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCT 673

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
                 ::::::::::||||||      ::::::          :::||||||||||||||||||
Db        674 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 733

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr------------------ 280
                 |||                        |||   ||||||
Db        734 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||:::|||||||||          ||||||   |||      |||
Db        773 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 832

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla------------------TyrAspHisArg 314
              |||      :::||||||   |||   :::                        :::   :::
Db        833 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
```

```
                       :::|||      :::   :::|||   ::::::   |||||||||
Db        893 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACT 952

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
              |||||||||:::::|||:::::|||          |||:::          :::   |||
Db        953 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1003

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
                     :::|||   :::|||:::::||||||||||||   :::||||||||||||
Db       1004 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1063

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
              :::|||||||||:::|||:::   :::|||          |||:::|||:::      |||
Db       1064 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1123

Qy        390 TyrLeuThrLys******AsnAspLeuArgGluGlnGlyGlnPro 404
                 |||   :::           :::::::::          |||
Db       1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168

RESULT 26
US-08-486-673B-62
; Sequence 62, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
;  APPLICANT: Sundelin, Johan
;  APPLICANT:  Scarborough, Robert M.
;  TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
;  FILE REFERENCE: 44481-5006-08-US
;  CURRENT APPLICATION NUMBER: US/08/486,673B
;  CURRENT FILING DATE:  1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/097,938
;  PRIOR FILING DATE: 1993-07-26
;  PRIOR APPLICATION NUMBER: PCT/US94/08536
;  PRIOR FILING DATE: 1994-07-26
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
;   LENGTH: 1414
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (50)..(1240)
;   OTHER INFORMATION: C140 receptor, cDNA and deduced protein sequences
US-08-486-673B-62

Alignment Scores:
Pred. No.:              7.59e-49        Length:         1414
Score:                  572.50          Matches:        140
Percent Similarity:     49.64%          Conservative:   66
Best Local Similarity:  33.73%          Mismatches:     152
Query Match:            26.54%          Indels:         57
DB:                     4               Gaps:           14

US-09-208-629F-6 (1-408) x US-08-486-673B-62 (1-1414)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
              |||   |||:::|||      :::           :::   |||   |||      |||
Db         41 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 100

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
              |||   |||                 |||   ||||||                ||||||
Db        101 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 139

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
                 :::|||           |||           |||
Db        140 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA-------------- 184

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
                          |||   :::   |||      |||   .|||:::
Db        185 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
```

```
Db        227 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 256

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127

Db        257 GCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGT 316

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147

Db        317 TTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 376

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166

Db        377 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 436

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186

Db        437 TTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 496

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206

Db        497 GTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 556

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226

Db        557 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 613

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246

Db        614 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCT 673

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266

Db        674 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 733

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr------------------ 280

Db        734 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300

Db        773 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 832

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla------------------TyrAspHisArg 314

Db        833 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334

Db        893 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACT 952

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354

Db        953 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1003

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371

Db       1004 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1063

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389

Db       1064 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1123

Qy        390 TyrLeuThrLys******AsnAspLeuArgGluGlnGlyGlnPro 404

Db       1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168

RESULT 27
US-09-053-866-1
; Sequence 1, Application US/09053866
; Patent No. 6111075
;  GENERAL INFORMATION:
;    APPLICANT:  Xu, Wenfeng
;    APPLICANT:  Presnell, Scott R.
;    APPLICANT:  Yee, David P.
;    APPLICANT:  Foster, Donald C.
;    TITLE OF INVENTION:  PROTEASE-ACTIVATED RECEPTOR
;    TITLE OF INVENTION:  PAR4 (ZCHEMR2)
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/053,866
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leith, Debra K
;      REGISTRATION NUMBER:  32,619
;      REFERENCE/DOCKET NUMBER:  98-10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6674
;      TELEFAX:  206-442-6678
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4895 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  176...1330
;      OTHER INFORMATION:
US-09-053-866-1

Alignment Scores:
Pred. No.:              4.33e-43        Length:         4895
Score:                  523.50          Matches:        121
Percent Similarity:     53.57%          Conservative:   59
Best Local Similarity:  36.01%          Mismatches:     145
Query Match:            24.27%          Indels:         12
DB:                     3               Gaps:           5

US-09-208-629F-6 (1-408) x US-09-053-866-1 (1-4895)

Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67

Db        225 CTGGCGGCACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGA-GGTGGTGAT 283

Qy         68 GluGluPhePro------PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85

Db        284 GACAGCACGCCCTCAATCCTGCCTGCCCCCCGCGGCTACCCAGGCCAAGTC--------- 334

Qy         86 LysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105

Db        335 ---------TGTGCCAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGGCA 382

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125

Db        383 CTGCTTCTGGGCTGGGTGCCCACCAGGCTGGTGCCCGCCCTCTATGGGCTGGTCCTGGTG 442

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
```

```
Db        443 GTGGGGCTGCCGGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACGCAGGCACCTCGGCTG 502

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
              :::|||:::          |||||||||   ||||||   |||       :::   |||
Db        503 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCCTGCTGGCCCTGGCGCTG 562

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
              |||   :::|||||||||||||||   |||       |||   |||||||||       |||
Db        563 CCCCCGCGGATCGCCTACCACCTGCGTGGCCAGCGCTGGCCCTTCGGGGAGGCCGCCTGC 622

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205
              |||       |||          ||||||:::||||||   |||:::|||||||||||||
Db        623 CGCCTGGCCACGGCCGCACTCTATGGTCACATGTATGGCTCAGTGCTGCTGCTGGCCGCC 682

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
              :::|||::::::|||||||||||:::|||||||||          |||   |||
Db        683 GTCAGCCTGGATCGCTACCTGGCCCTGGTGCACCCGCTGCGGGCCCGCGCCCTGCGTGGC 742

Qy        226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
              ||||||      |||          |||                           |||
Db        743 CGGCGCCTGGCCCTTGGACTCTGCATGGCTGCTTGGCTCATGGCGGCCGCCCTGGCACTG 802

Qy        246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
              |||       :::   :::|||   :::  |||   :::   |||         ||||||
Db        803 CCCCTGACACTGCAGCGGCAGACCTTCCGGCTGGCGCGCTCCGATCGCGTGCTCTGCCAT 862

Qy        266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
              |||                  :::|||   :::|||         |||       |||
Db        863 GACGCGCTGCCCCTGGACGCACAGGCCTCCCACTGGCAA---CCGGCCTTCACCTGCCTG 919

Qy        286 AlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305
              |||       |||      :::|||   :::   ::::::   ||||||   |||   :::
Db        920 GCGCTGTTGGGCTGTTTCCTGCCCCTGCTGGCCATGCTGCTGTGCTACGGGGCCACCCTG 979

Qy        306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325
              ||||||   |||          |||::  .      ::::::   :::   ::::::
Db        980 CACACGCTGGCGGCCAGCGGCCGGCGCTACGGCCACGCGCTGAGGCTGACCGCAGTGGTG 1039

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345
              |||                |||   |||||||||:::::::|||:::::::|||::::::
Db       1040 CTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGCAACCTGCTGCTGCTGCTGCATTACTCG 1099

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365
              :::          :::          ||||||      |||:::   :::|||   |||
Db       1100 GACCCGAGCCCCAGCGCCTGGGGCAACCTCTATGGTGCCTACGTGCCCAGCCTGGCGCTG 1159

Qy        366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
                 :::||||||||||||:::||||||||:::|||:::   :::|||
Db       1160 AGCACCCTCAACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1207

RESULT 28
US-09-479-130-1
; Sequence 1, Application US/09479130
; Patent No. 6436400
;  GENERAL INFORMATION:
;    APPLICANT:  Xu, Wenfeng
;    APPLICANT:  Presnell, Scott R.
;    APPLICANT:  Yee, David P.
;    APPLICANT:  Foster, Donald C.
;    TITLE OF INVENTION:  PROTEASE-ACTIVATED RECEPTOR
;    TITLE OF INVENTION:  PAR4 (ZCHEMR2)
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/479,130
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leith, Debra K
;      REGISTRATION NUMBER:  32,619
;      REFERENCE/DOCKET NUMBER:  98-10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6674
;      TELEFAX:  206-442-6678
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4895 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  176...1330
;      OTHER INFORMATION:
US-09-479-130-1

Alignment Scores:
Pred. No.:                4.33e-43        Length:          4895
Score:                    523.50          Matches:         121
Percent Similarity:       53.57%          Conservative:    59
Best Local Similarity:    36.01%          Mismatches:      145
Query Match:              24.27%          Indels:          12
DB:                       4               Gaps:            5

US-09-208-629F-6 (1-408) x US-09-479-130-1 (1-4895)

Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
              ||||||   |||      |||      |||   |||   |||||||||
Db        225 CTGGCGGCACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGA-GGTGGTGAT 283

Qy         68 GluGluPhePro------PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85
              :::      |||          |||      |||:::   |||
Db        284 GACAGCACGCCCTCAATCCTGCCTGCCCCCCGCGGCTACCCAGGCCAAGTC--------- 334

Qy         86 LysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105
                       |||      :::|||   :::   |||   :::   :::::::::
Db        335 ---------TGTGCCAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGGCA 382

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
                 |||         :::   |||:::|||:::||||||:::|||.  ||||||   |||
Db        383 CTGCTTCTGGGCTGGGTGCCCACCAGGCTGGTGCCCGCCCTCTATGGGCTGGTCCTGGTG 442

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
              ||||||:::|||||||||   :::   ||||||:::|||      :::         :::
Db        443 GTGGGGCTGCCGGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACGCAGGCACCTCGGCTG 502

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
              :::|||:::          |||||||||   ||||||   |||       :::   |||
Db        503 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCCTGCTGGCCCTGGCGCTG 562

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
              |||   :::|||||||||||||||   |||       |||   |||||||||       |||
Db        563 CCCCCGCGGATCGCCTACCACCTGCGTGGCCAGCGCTGGCCCTTCGGGGAGGCCGCCTGC 622

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205
              |||       |||          ||||||:::||||||   |||:::|||||||||||||
Db        623 CGCCTGGCCACGGCCGCACTCTATGGTCACATGTATGGCTCAGTGCTGCTGCTGGCCGCC 682

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
              :::|||::::::|||||||||||:::|||||||||          |||   |||
Db        683 GTCAGCCTGGATCGCTACCTGGCCCTGGTGCACCCGCTGCGGGCCCGCGCCCTGCGTGGC 742
```

```
Qy        226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
Db        743 CGGCGCCTGGCCCTTGGACTCTGCATGGCTGCTTGGCTCATGGCGGCCGCCCTGGCACTG 802

Qy        246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
Db        803 CCCCTGACACTGCAGCGGCAGACCTTCCGGCTGGCGCGCTCCGATCGCGTGCTCTGCCAT 862

Qy        266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
Db        863 GACGCGCTGCCCCTGGACGCACAGGCCTCCCACTGGCAA---CCGGCCTTCACCTGCCTG 919

Qy        286 AlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305
Db        920 GCGCTGTTGGGCTGTTTCCTGCCCCTGCTGGCCATGCTGCTGTGCTACGGGGCCACCCTG 979

Qy        306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325
Db        980 CACACGCTGGCGGCCAGCGGCCGGCGCTACGGCCACGCGCTGAGGCTGACCGCAGTGGTG 1039

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345
Db       1040 CTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGCAACCTGCTGCTGCTGCTGCATTACTCG 1099

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365
Db       1100 GACCCGAGCCCCAGCGCCTGGGGCAACCTCTATGGTGCCTACGTGCCCAGCCTGGCGCTG 1159

Qy        366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db       1160 AGCACCCTCAACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1207

RESULT 29
US-07-657-769B-68
; Sequence 68, Application US/07657769B
; Patent No. 5256766
;  GENERAL INFORMATION:
;    APPLICANT:  COUGHLIN, SHAUN R.
;    TITLE OF INVENTION:  RECOMBINANT RECEPTOR AND RELATED
;    TITLE OF INVENTION:  PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  IRELL & MANELLA
;      STREET:  545 MIDDLEFIELD ROAD, SUITE 200
;      CITY:  MENLO PARK
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/657,769B
;      FILING DATE:  19910219
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  2000-0502.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-327-7250
;      TELEFAX:  415-327-2951
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3480 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  225..1503
US-07-657-769B-68

Alignment Scores:
Pred. No.:                2.96e-43      Length:          3480
Score:                    523.00        Matches:         127
Percent Similarity:       52.59%        Conservative:    76
Best Local Similarity:    32.90%        Mismatches:      153
Query Match:              24.25%        Indels:          30
DB:                       1             Gaps:            11

US-09-208-629F-6 (1-408) x US-07-657-769B-68 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----------IleThrValLysIleLys 88
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
```

```
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              |||||||||   |||:::                  :::       |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::            |||      :::      ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||::::::||||||   |||::::::       :::   :::   |||:::
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::||||||
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1332 ATTTACTATTACGCTTCC 1349
```

```
RESULT 30
US-07-789-184-219
; Sequence 219, Application US/07789184
; Patent No. 5688768
;  GENERAL INFORMATION:
;    APPLICANT:  COUGHLIN, SHAUN R.
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT THROMBIN RECEPTOR AND
;    TITLE OF INVENTION:  RELATED PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  223
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/789,184
;      FILING DATE:  19911107
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22000-20502.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  34-0154
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3480 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  225..1499
US-07-789-184-219

Alignment Scores:
Pred. No.:              2.96e-43        Length:         3480
Score:                  523.00          Matches:        127
Percent Similarity:     52.59%          Conservative:   76
Best Local Similarity:  32.90%          Mismatches:     153
Query Match:            24.25%          Indels:         30
```

```
DB:                  1              Gaps:          11

US-09-208-629F-6 (1-408) x US-07-789-184-219 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||       |||:::       |||   ||||||
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||       :::                  ||||||                 |||
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||     ::::::|||            ||||||      :::|||   |||
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----------IleThrValLysIleLys 88
                    :::      :::|||   |||           :::::::::
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
              |||   :::                  :::                ||||||||||
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||      :::|||::::::|||      ||||||||||   :::
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||      |||:::   :::::::::|||      :::   :::
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::      :::||||||   ||||||   ||||||         ||||||
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||:::::|||   ||||||      |||||||||
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                 |||      ||||||   |||||||||   ||||||||||||:::      |||
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||           :::::|||   |||
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
              :::   ::::::          ||||||      :::||||||
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ::::::|||     :::  :::|||      :::          :::|||||||||
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                  :::      |||||||||
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::          |||   :::|||   ::::::          ||||||   :::
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              |||||||||   |||:::                  :::       |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::            |||      :::      ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
```

```
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375

Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381

Db       1332 ATTTACTATTACGCTTCC 1349

RESULT 31
US-08-475-263-219
; Sequence 219, Application US/08475263
; Patent No. 5759994
;  GENERAL INFORMATION:
;    APPLICANT:  COUGHLIN, SHAUN R.
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT THROMBIN RECEPTOR AND
;    TITLE OF INVENTION:  RELATED PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  223
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Ave., NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,263
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22000-20502.03
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3480 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  225..1499
US-08-475-263-219

Alignment Scores:
Pred. No.:              2.96e-43        Length:         3480
Score:                  523.00          Matches:        127
Percent Similarity:     52.59%          Conservative:   76
Best Local Similarity:  32.90%          Mismatches:     153
Query Match:            24.25%          Indels:         30
DB:                     1               Gaps:           11

US-09-208-629F-6 (1-408) x US-08-475-263-219 (1-3480)

Qy          14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33

Db         228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy          34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53

Db         288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT----------GCCACC 335

Qy          54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluProPhe 72

Db         336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy          73 SerAlaLeuGluGlyTrpThrGlyAlaThr----------IleThrValLysIleLys 88

Db         396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy          89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108

Db         456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy         109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128

Db         516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy         129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146

Db         576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy         147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166

Db         633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy         167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186

Db         693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy         187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206

Db         753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy         207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223

Db         813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy         224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243

Db         873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy         244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263

Db         924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy         264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283

Db         984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy         284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303

Db        1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy         304 IleIleArgThrLeuAsn----------AlaTyrAspHisArgTrpLeu 316

Db        1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy         317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336

Db        1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy         337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355

Db        1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy         356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375

Db        1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy         376 LeuTyrPheLeuMetSer 381

Db        1332 ATTTACTATTACGCTTCC 1349
```

```
RESULT 32
US-08-485-886-219
; Sequence 219, Application US/08485886
; Patent No. 5798248
;  GENERAL INFORMATION:
;    APPLICANT:  COUGHLIN, SHAUN R.
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT THROMBIN RECEPTOR AND
;    TITLE OF INVENTION:  RELATED PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  223
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,886
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/789,184
;      FILING DATE:  07-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22000-20502.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  34-0154
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3480 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  225..1499
US-08-485-886-219

Alignment Scores:
Pred. No.:              2.96e-43        Length:         3480
Score:                  523.00          Matches:        127
Percent Similarity:     52.59%          Conservative:   76
Best Local Similarity:  32.90%          Mismatches:     153
Query Match:            24.25%          Indels:         30
DB:                     1               Gaps:           11

US-09-208-629F-6 (1-408) x US-08-485-886-219 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||     |||:::     |||   ||||||
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||     :::                 ||||||               |||
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||     ::::::|||           ||||||     :::|||   |||
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr----------IleThrValLysIleLys 88
                 :::     :::|||   |||          :::::::::
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
              |||   :::                 :::                 ||||||||||
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||     |||      :::|||::::::|||      ||||||||||   :::
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||      |||:::   :::::::::|||     :::   :::
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::     :::||||||   ||||||   ||||||         ||||||
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||:::::|||   ||||||      |||||||||
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                 |||      ||||||   |||||||||   ||||||||||||:::      |||
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||            :::::::|||   |||
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
                 :::   ::::::         ||||||      :::|||||||
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              :::::::|||      :::   :::|||          :::          :::|||||||||
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                  :::      |||||||||
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::         |||   :::|||   ::::::      ||||||   :::
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              |||||||||   |||:::                 :::         |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::                 |||      :::      ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||::::::||||||   |||::::::         :::   :::   |||:::
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::||||||
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1332 ATTTACTATTACGCTTCC 1349

RESULT 33
US-08-477-362-219
; Sequence 219, Application US/08477362
; Patent No. 5849507
;  GENERAL INFORMATION:
;    APPLICANT:  COUGHLIN, SHAUN R.
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT THROMBIN RECEPTOR AND
```

```
;    TITLE OF INVENTION:  RELATED PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  223
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,362
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/789,184
;      FILING DATE:  07-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22000-20502.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  34-0154
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3480 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  225..1499
US-08-477-362-219

Alignment Scores:
Pred. No.:                2.96e-43        Length:          3480
Score:                    523.00          Matches:         127
Percent Similarity:       52.59%          Conservative:    76
Best Local Similarity:    32.90%          Mismatches:      153
Query Match:              24.25%          Indels:          30
DB:                       2               Gaps:            11

US-09-208-629F-6 (1-408) x US-08-477-362-219 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||       |||:::      |||   ||||||
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||       :::                  ||||||                 |||
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT------------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||       ::::::|||             ||||||      :::|||   |||
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr------------IleThrValLysIleLys 88
                    :::      :::|||   |||             :::::::::
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
                 |||   :::                  :::                ||||||||||
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||      :::|||::::::|||      ||||||||||||   :::
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||       |||:::   :::::::::|||       :::   :::
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                    |||:::      :::||||||   ||||||   ||||||         ||||||
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||::::::|||   ||||||      |||||||||
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                    |||       ||||||   |||||||||   ||||||||||||:::       |||
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||            ::::::|||   |||
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
                 :::   ::::::         ||||||      :::|||||
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ::::::|||       :::   :::|||         :::         :::|||||||||
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                    :::      |||||||||
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::         |||   :::|||   ::::::         ||||||   :::
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              |||||||||  |||:::                     :::         |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::          |||      :::      ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||::::::||||||   |||::::::      :::   :::   |||:::
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::||||||
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1332 ATTTACTATTACGCTTCC 1349

RESULT 34
US-08-477-134-219
; Sequence 219, Application US/08477134
; Patent No. 5856448
;  GENERAL INFORMATION:
;    APPLICANT:  COUGHLIN, SHAUN R.
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT THROMBIN RECEPTOR AND
;    TITLE OF INVENTION:  RELATED PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  223
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
```

```
;       ZIP:  94304-1018
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/477,134
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/789,184
;       FILING DATE:  07-NOV-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  MURASHIGE, KATE H.
;       REGISTRATION NUMBER:  29,959
;       REFERENCE/DOCKET NUMBER:  22000-20502.20
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 813-5600
;       TELEFAX:  (415) 494-0792
;       TELEX:  34-0154
;   INFORMATION FOR SEQ ID NO:  219:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3480 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  225..1499
US-08-477-134-219

Alignment Scores:
Pred. No.:                2.96e-43        Length:          3480
Score:                    523.00          Matches:         127
Percent Similarity:       52.59%          Conservative:    76
Best Local Similarity:    32.90%          Mismatches:      153
Query Match:              24.25%          Indels:          30
DB:                       2               Gaps:            11

US-09-208-629F-6 (1-408) x US-08-477-134-219 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||      |||:::      |||   ||||||
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||      :::                    ||||||                 |||
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||      ::::::|||           ||||||      :::|||   |||
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr----------IleThrValLysIleLys 88
                 :::      :::|||   |||            :::::::::
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
              |||   :::                :::                ||||||||||||
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||      :::|||:::::::|||      ||||||||||||   :::
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||      |||:::   :::::::::|||      :::   :::
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::      :::||||||   ||||||   ||||||         ||||||
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||::::::|||   ||||||      |||||||||
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                    |||      ||||||   |||||||||   ||||||||||||:::      |||
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||              :::::::|||   |||
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
                 :::   ::::::          ||||||      :::||||||
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ::::::|||      :::   :::|||      :::      :::||||||||||
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                  :::      |||||||||
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::          |||   :::|||   ::::::          ||||||   :::
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              |||||||||·  |||:::                  :::      |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::                |||      :::      ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||::::::||||||   |||::::::          :::   :::   |||:::
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::||||||
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1332 ATTTACTATTACGCTTCC 1349

RESULT 35
US-08-473-489A-219
; Sequence 219, Application US/08473489A
; Patent No. 6024936
;  GENERAL INFORMATION:
;    APPLICANT:  COUGHLIN, SHAUN R.
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT THROMBIN RECEPTOR AND
;    TITLE OF INVENTION:  RELATED PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  223
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/473,489A
```

```
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/789,184
;      FILING DATE:  1991-11-07
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22000-20502.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  34-0154
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3480 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  225..1499
US-08-473-489A-219

Alignment Scores:
Pred. No.:              2.96e-43        Length:         3480
Score:                  523.00          Matches:        127
Percent Similarity:     52.59%          Conservative:   76
Best Local Similarity:  32.90%          Mismatches:     153
Query Match:            24.25%          Indels:         30
DB:                     3               Gaps:           11

US-09-208-629F-6 (1-408) x US-08-473-489A-219 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   ::   |||     |||:::     |||   ||||||
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
              |||      :::                    ||||||                  |||
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||      ::::::|||              ||||||     :::|||   |||
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----------IleThrValLysIleLys 88
                 :::      :::|||   |||              :::::::::
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
              |||   :::                    :::                 ||||||||||||
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||     :::|||::::::|||      ||||||||||||   :::
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||      |||:::   :::::::::|||      :::   :::
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::      :::||||||   ||||||   ||||||          ||||||
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||::::::|||   ||||||      |||||||||
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                 |||      ||||||   |||||||||   ||||||||||:::      |||
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||            ::::::|||   |||
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
                 :::   ::::::          ||||||      :::|||||
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ::::::|||     :::   :::|||         :::         :::|||||||||
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                  :::      |||||||||
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::          |||   :::|||   ::::::        |||||   :::
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              |||||||||   |||:::                        :::         |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::              |||      :::      ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||::::::||||||   |||::::::      :::   :::   |||:::
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::|||||||
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1332 ATTTACTATTACGCTTCC 1349

RESULT 36
US-08-485-695-219
; Sequence 219, Application US/08485695
; Patent No. 6124101
;  GENERAL INFORMATION:
;    APPLICANT:  COUGHLIN, SHAUN R.
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT THROMBIN RECEPTOR AND
;    TITLE OF INVENTION:  RELATED PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  223
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,695
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/789,184
;      FILING DATE:  07-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
```

```
;      REFERENCE/DOCKET NUMBER:  22000-20502.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  34-0154
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3480 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  225..1499
US-08-485-695-219

Alignment Scores:
Pred. No.:              2.96e-43        Length:         3480
Score:                  523.00          Matches:        127
Percent Similarity:     52.59%          Conservative:   76
Best Local Similarity:  32.90%          Mismatches:     153
Query Match:            24.25%          Indels:         30
DB:                     3               Gaps:           11

US-09-208-629F-6 (1-408) x US-08-485-695-219 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||     |||:::     |||   ||||||
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||     :::                   ||||||                 |||
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||      ::::::|||           ||||||     :::|||   |||
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----------IleThrValLysIleLys 88
                 :::     :::|||   |||               :::::::::
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
              |||   :::                 :::                 ||||||||||||
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||      :::|||:::::::|||      ||||||||||||   :::
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||      |||:::   :::::::::|||      :::   :::
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::     :::||||||   ||||||   ||||||         ||||||
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||:::::::|||   ||||||      |||||||||
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                 |||      ||||||   |||||||||   ||||||||||||:::      |||
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||            ::::::|||   |||
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
                 :::   ::::::         ||||||      :::|||||
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ::::::|||      :::   :::|||         :::         :::|||||||||
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                 :::         ||||||||
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::         |||   :::|||   :::::         ||||||   :::
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              ||||||||   |||:::                     :::         |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::         |||     :::     ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||::::::||||||   |||::::::         :::   :::   |||:::
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::||||||
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1332 ATTTACTATTACGCTTCC 1349

RESULT 37
US-08-018-760-219
; Sequence 219, Application US/08018760
; Patent No. 6197541
;  GENERAL INFORMATION:
;    APPLICANT:  COUGHLIN, SHAUN R.
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  RECOMBINANT THROMBIN RECEPTOR AND
;    TITLE OF INVENTION:  RELATED PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  223
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/018,760
;      FILING DATE:  17-FEB-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/789,184
;      FILING DATE:  07-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22000-20502.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  34-0154
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3480 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  225..1499
US-08-018-760-219

Alignment Scores:
Pred. No.:                2.96e-43      Length:          3480
Score:                    523.00        Matches:         127
Percent Similarity:       52.59%        Conservative:    76
Best Local Similarity:    32.90%        Mismatches:      153
Query Match:              24.25%        Indels:          30
DB:                       4             Gaps:            11

US-09-208-629F-6 (1-408) x US-08-018-760-219 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||      |||:::      |||   ||||||
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||      :::                  ||||||                |||
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT------------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||      ::::::|||           ||||||      :::|||   |||
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr------------IleThrValLysIleLys 88
                    :::      :::|||   |||            :::::::::
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
                 |||   :::                  :::                ||||||||||||
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||      :::|||::::::|||      ||||||||||   :::
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||      |||:::   :::::::::|||      :::   :::
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::      :::||||||   ||||||   ||||||          ||||||
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||::::::|||   ||||||      ||||||||
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                 |||      ||||||   ||||||||   ||||||||||:::      |||
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||            ::::::|||   |||
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
                 :::   ::::::         ||||||      :::||||||
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ::::::|||      :::   :::|||          :::          :::|||||||||
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                  :::      ||||||||
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::          |||   :::|||   ::::::          ||||||   :::
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              |||||||||   |||:::                  :::          |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::            |||      :::      ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||::::::||||||   |||::::::         :::   :::   |||:::
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::||||||
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1332 ATTTACTATTACGCTTCC 1349

RESULT 38
US-08-313-553-12
; Sequence 12, Application US/08313553
; Patent No. 5641650
;  GENERAL INFORMATION:
;    APPLICANT:  TURNER, George J.
;    APPLICANT:  BETLACH, Mary C.
;    TITLE OF INVENTION:  EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
;    TITLE OF INVENTION:  IN HALOBACTERIA
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Walter H. Dreger
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/313,553
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/038,662
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Walter H.
;      REGISTRATION NUMBER:  24,190
;      REFERENCE/DOCKET NUMBER:  A-57669/WHD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1764 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  436..462
;      OTHER INFORMATION:  /note= "Sequence encoding
; Patent No. 5641650
```

```
;      OTHER INFORMATION:  polyaspartic acid."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  463..465
;      OTHER INFORMATION:  /note= "Codon encoding the
;      OTHER INFORMATION:  N-terminal amino acid of the human thrombin
;      OTHER INFORMATION:  receptor protein."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1630..1632
;      OTHER INFORMATION:  /note= "Codon encoding the
;      OTHER INFORMATION:  C-terminal amino acid of the human thrombin
;      OTHER INFORMATION:  receptor protein. "
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  1633..1650
;      OTHER INFORMATION:  /note= "Sequence encoding
; Patent No. 5641650
;      OTHER INFORMATION:  polyhistidine."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  648..656
;      OTHER INFORMATION:  /note= "Deleted AlwNI restriction
;      OTHER INFORMATION:  site."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  893..898
;      OTHER INFORMATION:  /note= "Deleted PstI restriction
;      OTHER INFORMATION:  site."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1301..1309
;      OTHER INFORMATION:  /note= "Deleted AlwNI restriction
;      OTHER INFORMATION:  site."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1394..1402
;      OTHER INFORMATION:  /note= "Deleted AlwNI restriction
;      OTHER INFORMATION:  site."
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  374
;      OTHER INFORMATION:  /note= "RNA start site."
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(1671, "")
;      OTHER INFORMATION:  /note= "C to T mutation removes
;      OTHER INFORMATION:  AlwNI site."
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  376..1650
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  376..414
;      OTHER INFORMATION:  /note= "Bacteriorhodopsin
;      OTHER INFORMATION:  pre-sequence."
;    FEATURE:
;      NAME/KEY:  terminator
;      LOCATION:  1672..1674
;      OTHER INFORMATION:  /note= "Bacteriorhodopsin stop
;      OTHER INFORMATION:  codon."
;    FEATURE:
;      NAME/KEY:  terminator
;      LOCATION:  1651..1653
;      OTHER INFORMATION:  /note= "Thrombin stop codon."
US-08-313-553-12

Alignment Scores:
Pred. No.:              5.05e-43        Length:         1764
Score:                  516.50          Matches:        123
Percent Similarity:     53.05%          Conservative:   77
Best Local Similarity:  32.63%          Mismatches:     144
Query Match:            23.95%          Indels:         33
```

```
DB:                    1              Gaps:          11

US-09-208-629F-6 (1-408) x US-08-313-553-12 (1-1764)

Qy         30 LeuLeuLeuLeuLeuProThrPheCysGln--------------------SerGlyMet 42
              :::|||   ||||||||||||      :::                        :::
Db        376 ATGTTGGAGTTATTGCCAACAGCAGTGGAGGGGGTATCGCAGGCCCAGATCCAGGCGCTG 435

Qy         43 GluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAla 62
                :::         ::::::          ||||||      :::::::|||
Db        436 GACTACAAGGACGATGATGACGTCGACGCCACCTTAGATCCCCGGTCATTTCTTCTCAGG 495

Qy         63 ProProAsn---SerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
                 ||||||      :::|||   |||                 :::      :::|||
Db        496 AACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGGTTA 555

Qy         82 Thr------------IleThrValLysIleLysCysProGluGluSerAlaSerHisLeu 97
              |||            :::::::::              |||   :::
Db        556 ACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTTCAAAAACAACTTCCTGCA 615

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
              :::                ||||||||||||||||||      |||      :::|||
Db        616 TTCATCTCAGAAGATGCCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCA 675

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsn------AlaValThrLeu 135
              ::::::|||      ||||||||||   :::|||   |||      |||:::   :::
Db        676 TCTGTGTACACCGGAGTGTTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTG 735

Qy        136 TrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAla 155
              ::::::|||      :::  :::              |||:::      :::||||||
Db        736 TTCATCCTGAAAATGAAGGTCAAGAAG---CCGGCGGTGGTGTACATGCTGCACCTGGCC 792

Qy        156 IleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGly 175
               ||||||   ||||||          ||||||||||||||:::|||:::   :::|||
Db        793 ACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCCTTTAAGATCAGCTATTACTTTTCCGGC 852

Qy        176 AsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsn 195
              ::::::|||   ||||||      |||||||||      |||      ||||||   |||
Db        853 AGTGATTGGCAGTTTGGGTCTGAATTGTGTCGCTTCGTCACTGCAGCATTTTACTGTAAC 912

Qy        196 MetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleVal 215
              ||||||   ||||||||||||:::      |||||||||:::|||:::||||||:::|||
Db        913 ATGTACGCCTCTATCTTGCTCATGACAGTCATAAGCATTGACCGGTTTCTGGCTGTGGTG 972

Qy        216 HisPro---------PheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThr 232
              :::|||           ::::::|||  |||   :::   ::::::          |||
Db        973 TATCCCATGCAGTCCCTCTCCTGGCGTACTCTGGGAAGGGCTTCCTTC---------ACT 1023

Qy        233 CysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGln 252
              |||      :::|||||||                ::::::|||      :::   :::|||
Db       1024 TGTCTGGCCATCTGGGCTTTGGCCATCGCAGGGGTAGTGCCTCTCGTCCTCAAGGAGCAA 1083

Qy        253 GluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGlu 272
                       :::      :::||||||||||||||||||||||    |||
Db       1084 ACCATCCAGGTGCCCGGGCTCAACATCACTACCTGTCATGATGTGCTCAATGAAACCCTG 1143

Qy        273 SerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
                          :::      |||||||||   :::   :::          |||   :::
Db       1144 CTCGAAGGC---TACTATGCCTACTACTTCTCAGCCTTCTCTGCTGTCTTCTTTTTTGTG 1200

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsn--------- 309
              |||   ::::::          ||||||   :::|||||||||    |||:::
Db       1201 CCGCTGATCATTTCCACGGTCTGTTATGTGTCTATCATTCGATGTCTTAGCTCTTCCGCA 1260

Qy        310 ------------AlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325
                         :::      |||   |||:::                |||      :::
Db       1261 GTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTGTTC---------CTGTCAGCTGCTGTT 1311

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345
                 ||||||   |||||||||   |||:::|||::::::||||||   |||::::::
Db       1312 TTCTGCATCTTCATCATTTGCTTCGGACCCACAAACGTCCTCCTGATTGCGCATTACTCA 1371

Qy        346 AsnTyr---TyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCys 364
```

```
                    :::   :::   |||:::      ||||||   ||||||:::   :::|||
Db       1372 TTCCTTTCTCACACTTCCACCACAGAGGCTGCCTACTTTGCCTACCTCCTCTGTGTCTGT 1431

Qy        365 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
              :::   |||::::::||||||:::||||||   :::|||:::      |||
Db       1432 GTCAGCAGCATAAGCTCGTGCATCGACCCCCTAATTTACTATTACGCTTCC 1482

RESULT 39
US-08-767-993-12
; Sequence 12, Application US/08767993
; Patent No. 6010885
;  GENERAL INFORMATION:
;    APPLICANT:  TURNER, George J.
;    APPLICANT:  BETLACH, Mary C.
;    TITLE OF INVENTION:  EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
;    TITLE OF INVENTION:  IN HALOBACTERIA
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Walter H. Dreger
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/767,993
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Walter H.
;      REGISTRATION NUMBER:  24,190
;      REFERENCE/DOCKET NUMBER:  A-57669/WHD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1764 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  436..462
;      OTHER INFORMATION:  /note= "Sequence encoding
; Patent No. 6010885
;      OTHER INFORMATION:  polyaspartic acid."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  463..465
;      OTHER INFORMATION:  /note= "Codon encoding the
;      OTHER INFORMATION:  N-terminal amino acid of the human thrombin
;      OTHER INFORMATION:  receptor protein."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1630..1632
;      OTHER INFORMATION:  /note= "Codon encoding the
;      OTHER INFORMATION:  C-terminal amino acid of the human thrombin
;      OTHER INFORMATION:  receptor protein. "
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  1633..1650
;      OTHER INFORMATION:  /note= "Sequence encoding
; Patent No. 6010885
;      OTHER INFORMATION:  polyhistidine."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  648..656
;      OTHER INFORMATION:  /note= "Deleted AlwNI restriction
;      OTHER INFORMATION:  site."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  893..898
;      OTHER INFORMATION:  /note= "Deleted PstI restriction
;      OTHER INFORMATION:  site."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1301..1309
;      OTHER INFORMATION:  /note= "Deleted AlwNI restriction
;      OTHER INFORMATION:  site."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1394..1402
;      OTHER INFORMATION:  /note= "Deleted AlwNI restriction
;      OTHER INFORMATION:  site."
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  374
;      OTHER INFORMATION:  /note= "RNA start site."
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(1671, "")
;      OTHER INFORMATION:  /note= "C to T mutation removes
;      OTHER INFORMATION:  AlwNI site."
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  376..1650
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  376..414
;      OTHER INFORMATION:  /note= "Bacteriorhodopsin
;      OTHER INFORMATION:  pre-sequence."
;    FEATURE:
;      NAME/KEY:  terminator
;      LOCATION:  1672..1674
;      OTHER INFORMATION:  /note= "Bacteriorhodopsin stop
;      OTHER INFORMATION:  codon."
;    FEATURE:
;      NAME/KEY:  terminator
;      LOCATION:  1651..1653
;      OTHER INFORMATION:  /note= "Thrombin stop codon."
US-08-767-993-12

Alignment Scores:
Pred. No.:              5.05e-43        Length:         1764
Score:                  516.50          Matches:        123
Percent Similarity:     53.05%          Conservative:   77
Best Local Similarity:  32.63%          Mismatches:     144
Query Match:            23.95%          Indels:         33
DB:                     3               Gaps:           11

US-09-208-629F-6 (1-408) x US-08-767-993-12 (1-1764)

Qy         30 LeuLeuLeuLeuLeuProThrPheCysGln----------------------SerGlyMet 42
              :::|||   ||||||||||||      :::                       :::
Db        376 ATGTTGGAGTTATTGCCAACAGCAGTGGAGGGGGTATCGCAGGCCCAGATCCAGGCGCTG 435

Qy         43 GluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAla 62
              :::         ::::::            ||||||      ::::::|||
Db        436 GACTACAAGGACGATGATGACGTCGACGCCACCTTAGATCCCCGGTCATTTCTTCTCAGG 495

Qy         63 ProProAsn---SerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
                ||||||       :::|||   |||               :::      :::|||
Db        496 AACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGGTTA 555

Qy         82 Thr------------IleThrValLysIleLysCysProGluGluSerAlaSerHisLeu 97
              |||            :::::::::          |||   :::
Db        556 ACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTTCAAAAACAACTTCCTGCA 615

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
```

```
Db        616 TTCATCTCAGAAGATGCCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCA 675

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsn------AlaValThrLeu 135

Db        676 TCTGTGTACACCGGAGTGTTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTG 735

Qy        136 TrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAla 155

Db        736 TTCATCCTGAAAATGAAGGTCAAGAAG---CCGGCGGTGGTGTACATGCTGCACCTGGCC 792

Qy        156 IleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGly 175

Db        793 ACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCCTTTAAGATCAGCTATTACTTTTCCGGC 852

Qy        176 AsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsn 195

Db        853 AGTGATTGGCAGTTTGGGTCTGAATTGTGTCGCTTCGTCACTGCAGCATTTTACTGTAAC 912

Qy        196 MetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleVal 215

Db        913 ATGTACGCCTCTATCTTGCTCATGACAGTCATAAGCATTGACCGGTTTCTGGCTGTGGTG 972

Qy        216 HisPro---------PheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThr 232

Db        973 TATCCCATGCAGTCCCTCTCCTGGCGTACTCTGGGAAGGGCTTCCTTC---------ACT 1023

Qy        233 CysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGln 252

Db       1024 TGTCTGGCCATCTGGGCTTTGGCCATCGCAGGGGTAGTGCCTCTCGTCCTCAAGGAGCAA 1083

Qy        253 GluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGlu 272

Db       1084 ACCATCCAGGTGCCCGGGCTCAACATCACTACCTGTCATGATGTGCTCAATGAAACCCTG 1143

Qy        273 SerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292

Db       1144 CTCGAAGGC---TACTATGCCTACTACTTCTCAGCCTTCTCTGCTGTCTTCTTTTTTGTG 1200

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsn--------- 309

Db       1201 CCGCTGATCATTTCCACGGTCTGTTATGTGTCTATCATTCGATGTCTTAGCTCTTCCGCA 1260

Qy        310 ------------AlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325

Db       1261 GTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTGTTC---------CTGTCAGCTGCTGTT 1311

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345

Db       1312 TTCTGCATCTTCATCATTTGCTTCGGACCCACAAACGTCCTCCTGATTGCGCATTACTCA 1371

Qy        346 AsnTyr---TyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCys 364

Db       1372 TTCCTTTCTCACACTTCCACCACAGAGGCTGCCTACTTTGCCTACCTCCTCTGTGTCTGT 1431

Qy        365 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381

Db       1432 GTCAGCAGCATAAGCTCGTGCATCGACCCCCTAATTTACTATTACGCTTCC 1482

RESULT 40
US-08-383-750-3
; Sequence 3, Application US/08383750
; Patent No. 5744301
;  GENERAL INFORMATION:
;    APPLICANT:  Birkenbach, Mark
;    APPLICANT:  Kieff, Elliot
;    TITLE OF INVENTION:  Epstein Barr Virus Induced Genes
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue, N.W.,
;      STREET:  Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/383,750
;      FILING DATE:  Herewith
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fox, Samuel, L.
;      REGISTRATION NUMBER:  30,353
;      REFERENCE/DOCKET NUMBER:  0627.3300001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1643 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  34..1116
US-08-383-750-3

Alignment Scores:
Pred. No.:                2.61e-32        Length:          1643
Score:                    410.50          Matches:         99
Percent Similarity:       47.87%          Conservative:    58
Best Local Similarity:    30.18%          Mismatches:      114
Query Match:              19.03%          Indels:          57
DB:                       1               Gaps:            10

US-09-208-629F-6 (1-408) x US-08-383-750-3 (1-1643)

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125

Db        103 TATGCACATCACAGCACGGCCAGGATAGTAATGCCTCTGCATTACAGCCTCGTCTTCATC 162

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145

Db        163 ATTGGGCTCGTGGGAAACTTACTAGCCTTGGTCGTCATTGTTCAAAACAGGAAAAAAATC 222

Qy        146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164

Db        223 AACTCTACCACCCTCTATTCAACAAATTTGGTGATTTCTGATATACTTTTTACCACGGCT 282

Qy        165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184

Db        283 TTGCCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGAGAATCGGAGATGCCTTG 342

Qy        185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAla 204

Db        343 TGTAGGATAACTGCGCTAGTGTTTTACATCAACACATATGCAGGTGTGAACTTTATGACC 402

Qy        205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224

Db        403 TGCCTGAGTATTGACCGCTTCATTGCTGTGGTGCACCCTCTACGCTACAACAAGATAAAA 462

Qy        225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244

Db        463 AGGATTGAACATGCAAAAGGCGTGTGCATATTTGTCTGGATTCTAGTATTTGCTCAGACA 522

Qy        245 LeuProPhePheIle---------LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261

Db        523 CTCCCACTCCTCATCAACCCTATGTCAAAGCAGGAG----------GCTGAAAGG 567

Qy        262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer---ProPheGlnLeuTyr 280
```

```
                ||||||   :::      |||      |||:::   |||   |||:::   |||
Db        568 ATTACATGCATGGAGTATCCAAACTTTGAAGAAACTAAATCTCTTCCCTGGATTCTG--- 624

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
                 :::         |||   |||:::::::::|||   ::::::|||:::   |||
Db        625 ------CTTGGGGCATGTTTCATAGGATATGTACTTCCACTTATAATCATTCTCATCTGC 678

Qy        301 TyrAlaAlaIleIle---------------------------------------------- 305
                |||:::   |||
Db        679 TATTCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAAACCCACTCACTGAGAAA 738

Qy        306 --------------ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
                           :::   ||||||
Db        739 TCTGGTGTAAACAAAAAGGCTCTCAACACAATT-------------------------- 771

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
                 :::||||||:::|||:::|||   :::||||||   |||   ::::::   :::
Db        772 ------ATTCTTATTATTGTTGTGTTTGTTCTCTGTTTCACACCTTACCATGTTGCAATT 825

Qy        341 IleIleHis-----------------HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
                |||   |||                :::|||:::         :::
Db        826 ATTCAACATATGATTAAGAAGCTTCGTTTCTCTAATTTCCTGGAATGTAGCCAAAGACAT 885

Qy        355 LeuTyrPheIleTyrLeu---IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
               :::   |||   |||         :::||||||   :::   |||   |||:::|||
Db        886 TCGTTCCAGATTTCTCTGCACTTTACAGTATGCCTGATGAACTTCAATTGCTGCATGGAC 945

Qy        374 ProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys 393
                ||||||:::||||||      |||                 |||       |||
Db        946 CCTTTTATCTACTTCTTTGCATGTAAA-----------------GGGTATAAGAGAAAG 987

Qy        394 ******AsnAspLeuArgGluGln 401
                     |||:::   |||
Db        988 GTTATGAGGATGCTGAAACGGCAA 1011
```

Search completed: June 29, 2003, 10:17:56
Job time : 104.056 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:26:54 ; Search time 192.736 Seconds
(without alignments)
4767.217 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIKM.....AYLTKXXNDLREQGQPSQRT 408

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09208629/runat_24062003_101625_12912/app_query.fasta_1.1166
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629_@CGN_1_1_275_@runat_24062003_101625_12912 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2145	99.4	6546	21	AAA35311	Human adenosine re
2	2142	99.3	1224	19	AAV07374	Human protease-act
3	2142	99.3	1830	21	AAF21430	Human low adenosin
4	2142	99.3	1830	21	AAA35308	Human adenosine re
5	2142	99.3	6546	21	AAF21433	Human low adenosin
6	1947	90.3	1102	19	AAV07375	Human protease-act
7	1443.5	66.9	1224	19	AAV07372	Mouse protease-act
8	1443.5	66.9	2409	21	AAZ50774	Human protease act
9	959	44.5	1124	19	AAV07373	Mouse protease-act
10	606.5	28.1	2732	16	AAQ84559	Murine C140 recept
11	606.5	28.1	2732	17	AAT32038	Murine C140 recept
12	600.5	27.8	1477	16	AAQ84557	Murine C140 recept
13	600.5	27.8	1477	17	AAT32036	Murine C140 recept
14	586.5	27.2	1451	21	AAC60319	Human PAR-2 DNA.
15	585.5	27.1	1255	16	AAQ84558	Human C140 recepto
16	585.5	27.1	1255	17	AAT32037	Human C140 recepto
17	585.5	27.1	1289	21	AAZ50773	Human protease act
18	572.5	26.5	1414	16	AAQ84560	Human C140 recept
19	572.5	26.5	1414	17	AAT32039	Human C140 recepto
20	531.5	24.6	1361	20	AAX90979	Mouse protease-act
21	523.5	24.3	4895	20	AAZ32747	Human protease-act
22	523.5	24.3	4895	21	AAZ50775	Human protease act
23	523.5	24.3	4925	22	AAH43632	Human PAR4 coding
24	523	24.2	1534	20	AAX90981	Human protease-act
25	523	24.2	1534	21	AAF21431	Human low adenosin
26	523	24.2	1534	21	AAA35309	Human adenosine re
27	523	24.2	3472	20	AAZ32191	Human thrombin rec
28	523	24.2	3480	21	AAZ50771	Human thrombin rec
29	519	24.1	1278	24	ABK70889	Human cDNA encodin
30	516.5	23.9	1764	15	AAQ73590	Fragment of the hu
31	512	23.7	2910	18	AAT62461	ThR-GPA1 fusion ge
32	512	23.7	3182	21	AAF21432	Human low adenosin
33	512	23.7	3182	21	AAA35310	Human adenosine re
34	507.5	23.5	1209	24	ABK70887	Human cDNA encodin
35	504	23.4	3480	13	AAQ28568	Human thrombin rec
36	503.5	23.3	1116	24	ABK70888	Human cDNA encodin
37	496.5	23.0	1312	18	AAT93367	Mouse thrombin rec
38	488	22.6	2588	19	AAV40372	Mouse G-protein co
39	486.5	22.6	1080	19	AAV40373	Mouse G-protein co
40	466	21.6	1955	21	AAZ99588	cDNA encoding a hu
41	465	21.6	1192	21	AAA14828	cDNA encoding a hu
42	465	21.6	2137	22	AAH51006	Human nGPCR5 codin
43	461	21.4	1080	21	AAD01131	Human orphan G pro
44	461	21.4	1080	21	AAA46032	Human G protein co
45	460	21.3	1080	22	AAF57649	Human G-protein co

ALIGNMENTS

RESULT 1
AAA35311
ID AAA35311 standard; DNA; 6546 BP.
XX
AC AAA35311;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:185.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000..
XX
PF 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -
XX
PS Disclosure; Page 1331-1333; 1343pp; English.
XX
CC The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A²containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.
XX
SQ Sequence 6546 BP; 1603 A; 1692 C; 1500 G; 1750 T; 1 other;

Alignment Scores:

Pred. No.:	9.57e-188	Length:	6546
Score:	2145.00	Matches:	405
Percent Similarity:	99.26%	Conservative:	0
Best Local Similarity:	99.26%	Mismatches:	3
Query Match:	99.44%	Indels:	0
DB:	21	Gaps:	0

US-09-208-629F-6 (1-408) x AAA35311 (1-6546)

```
Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
              |||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db         88 TGCTCCATGATTTTACAGATTTCACAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 147

Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 207

Qy         41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 267

Qy         61 GlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 327

Qy         81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 387

Qy        101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 447

Qy        121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTCTTC 507

Qy        141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 AGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTTCTT 567

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTATTT 627

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 687

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 CTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTAC 747

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACAGTT 807

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 867

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 ATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTCTAT 927

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTACTGC 987

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 1047

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 GCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATTCTT 1107

Qy        341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATATCTC 1167

Qy        361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 ATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTCATG 1227

Qy        381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
              ||||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db       1228 TCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGAGAA 1287

Qy        401 GlnGlyGlnProSerGlnArgThr 408
              |||||||||||||||||||||||
Db       1288 CAAGGACAGCCATCACAGAGAACG 1311
```

RESULT 2
AAV07374
ID AAV07374 standard; cDNA; 1224 BP.
XX
AC AAV07374;
XX
DT 12-OCT-1998 (first entry)
XX
DE Human protease-activated receptor 3 (PAR3) cDNA.
XX
KW Protease-activated receptor 3; PAR3; thrombin receptor; human;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 58..1182
FT /*tag= a
XX
PN WO9818456-A1.
XX
PD 07-MAY-1998.
XX
PF 29-OCT-1997; 97WO-US19732.
XX
PR 30-OCT-1996; 96US-0742440.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Connolly A, Coughlin SR, Ishihara H;
XX
DR WPI; 1998-271905/24.
DR P-PSDB; AAW51406.
XX
PT DNA encoding protease-activated receptor 3 - for detection of
PT specific agonists and antagonists, potentially useful for treating
PT e.g. thrombosis, atherosclerosis, inflammation etc.
XX
PS Claim 3; Page 40-41; 74pp; English.
XX
CC This cDNA clone codes for mouse protease-activated receptor 3
CC (PAR3) (see AAW51406), a cell surface protein which is specifically
CC activated by thrombin or a thrombin agonist, thereby activating
CC signalling events such as phosphoinositide hydrolysis, calcium ion
CC efflux and platelet aggregation. A cDNA library was establised
CC from rat platelets and amplified (primers given in AAV07376-77). A
CC product was identified that was expected to encode a new G-protein
CC coupled receptor, related to PAR1 or 2. This was used to isolate
CC human and murine DNA by a combination of PCR and hybridisation
CC techniques. Murine cDNA (see AAV07372), murine genomic DNA (see
CC AAV07373), human cDNA and human genomic DNA (see AAV07375) sequences
CC are provided, as well as the murine (see AAW51405) and human PAR3
CC proteins. Also claimed are vectors, host cells and an assay
CC device. Host cells are used to screen compounds for their ability
CC to act as agonists or antagonists of the effects of thrombin-PAR3
CC interaction. Agonists are used to treat wounds, thrombosis,
CC atherosclerosis, restenosis, inflammation and other thrombin
CC activated disorders. Antagonists (see AAW51415-21) are used to
CC control blood coagulation and thereby to treat heart attack and
CC stroke. They also mediate inflammatory and proliferative responses
CC to injury as occur in wound healing, atherosclerosis, restenosis,
CC pulmonary inflammation (ARDS) and glomerulosclerosis.
XX
SQ Sequence 1224 BP; 303 A; 316 C; 221 G; 384 T; 0 other;

Alignment Scores:

Pred. No.:	2.08e-188	Length:	1224
Score:	2142.00	Matches:	405
Percent Similarity:	99.26%	Conservative:	0
Best Local Similarity:	99.26%	Mismatches:	3

Query Match:	99.30%	Indels:	0
DB:	19	Gaps:	0

US-09-208-629F-6 (1-408) x AAV07374 (1-1224)

```
Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||||
Db          1 TGCTCCATGATTTTACAGATTTCATAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 60

Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 120

Qy         41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 180

Qy         61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 240

Qy         81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 300

Qy        101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 360

Qy        121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTCTTC 420

Qy        141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTTCTT 480

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTATTT 540

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 600

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTAC 660

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACAGTT 720

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 780

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTCTAT 840

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTACTGC 900

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 960

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATTCTT 1020
```

```
Qy        341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATATCTC 1080

Qy        361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTCATG 1140

Qy        381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
              ||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db       1141 TCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGAGAA 1200

Qy        401 GlnGlyGlnProSerGlnArgThr 408
              |||||||||||||||||||||||
Db       1201 CAAGGACAGCCATCACAGAGAACG 1224
```

```
RESULT 3
AAF21430
ID   AAF21430 standard; DNA; 1830 BP.
XX
AC   AAF21430;
XX
DT   14-MAR-2001  (first entry)
XX
DE   Human low adenosine antisense oligonucleotide related sequence #2997.
XX
KW   Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW   human; airway disorder; bronchoconstriction; lung inflammation;
KW   surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW   immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW   respiratory obstruction; pulmonary obstruction; impeded respiration;
KW   surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW   respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW   pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW   chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200062736-A2.
XX
PD   26-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US08020.
XX
PR   06-APR-1999;   99US-0127958.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
PA   (NYCE/) NYCE J W.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-679539/66.
XX
PT   Low adenosine (A) content antisense oligonucleotides which do not
PT   trigger adenosine receptors during metabolism, useful e.g. for treating
PT   cancers and respiratory obstructions -
XX
PS   Disclosure; Page 1414-1415; 1592pp; English.
XX
CC   The present invention describes low adenosine (A) content antisense
CC   oligonucleotides and compositions (I) comprising them. In the antisense
CC   oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC   (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC   immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC   The antisense oligonucleotides and (I) can be used to down-regulate the
CC   expression and or activity of target polypeptides associated with
CC   lung/respiratory disorders and malignancies, such as stimulating and
CC   activating peptide factors and transmitters, transcription factors,
CC   immunoglobulins and antibodies, antibody receptors, cytokines and
CC   chemokines, endogenously produced specific and non-specific enzymes,
CC   binding proteins, adhesion molecules and their receptors, cytokine and
CC   chemokine receptors, adenosine receptors, bradykinin receptors, central
CC   nervous system (CNS) and peripheral nervous and non-nervous system
CC   receptors, CNS and peripheral nervous and non-nervous system peptide
CC   transmitters, defensins, growth factors, vasoactive peptides and
CC   receptors, binding proteins and malignancy associated proteins. The
CC   antisense oligonucleotides may be used in this way to treat disorders
CC   including respiratory obstruction (especially pulmonary obstruction
CC   and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC   and/or surfactant hypoproduction which are associated with a disease or
CC   condition selected from pulmonary vasoconstriction, inflammation,
CC   allergies, asthma, impeded respiration, respiratory distress syndrome
CC   (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC   hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC   pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC   and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC   fragments and antisense oligonucleotides used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 1830 BP; 473 A; 464 C; 337 G; 556 T; 0 other;
```

```
Alignment Scores:
Pred. No.:              3.49e-188      Length:          1830
Score:                  2142.00        Matches:         405
Percent Similarity:     99.26%         Conservative:    0
Best Local Similarity:  99.26%         Mismatches:      3
Query Match:            99.30%         Indels:          0
DB:                     21             Gaps:            0

US-09-208-629F-6 (1-408) x AAF21430 (1-1830)

Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
              |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db         88 TGCTCCATGATTTTACAGATTTCATAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 147

Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 207

Qy         41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 267

Qy         61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 327

Qy         81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 387

Qy        101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 447

Qy        121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTCTTC 507

Qy        141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 AGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTTCTT 567

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTATTT 627

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 687

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 CTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTAC 747
```

```
Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACAGTT 807

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 867

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 ATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTCTAT 927

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTACTGC 987

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 1047

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 GCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATTCTT 1107

Qy        341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATATCTC 1167

Qy        361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 ATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTCATG 1227

Qy        381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
              ||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db       1228 TCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGAGAA 1287

Qy        401 GlnGlyGlnProSerGlnArgThr 408
              |||||||||||||||||||||||||
Db       1288 CAAGGACAGCCATCACAGAGAACG 1311
```

RESULT 4
AAA35308
ID AAA35308 standard; DNA; 1830 BP.
XX
AC AAA35308;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:182.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -
XX
PS Disclosure; Page 1329; 1343pp; English.
XX
CC The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.
XX
SQ Sequence 1830 BP; 473 A; 464 C; 337 G; 556 T; 0 other;

Alignment Scores:

Pred. No.:	3.49e-188	Length:	1830
Score:	2142.00	Matches:	405
Percent Similarity:	99.26%	Conservative:	0
Best Local Similarity:	99.26%	Mismatches:	3
Query Match:	99.30%	Indels:	0
DB:	21	Gaps:	0

US-09-208-629F-6 (1-408) x AAA35308 (1-1830)

```
Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db         88 TGCTCCATGATTTTACAGATTTCATAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 147

Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 207

Qy         41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 267

Qy         61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 327

Qy         81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 387

Qy        101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 447

Qy        121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        448 CTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTCTTC 507

Qy        141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 AGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTTCTT 567

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTATTT 627

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 687

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 CTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTAC 747

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACAGTT 807

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 867

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 ATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTCTAT 927

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTACTGC 987

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 1047

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 GCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATTCTT 1107

Qy        341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATATCTC 1167

Qy        361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 ATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTCATG 1227

Qy        381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
              ||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db       1228 TCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGAGAA 1287

Qy        401 GlnGlyGlnProSerGlnArgThr 408
              |||||||||||||||||||||||
Db       1288 CAAGGACAGCCATCACAGAGAACG 1311
```

RESULT 5
AAF21433
ID AAF21433 standard; DNA; 6546 BP.
XX
AC AAF21433;
XX
DT 14-MAR-2001 (first entry)
XX
DE Human low adenosine antisense oligonucleotide related sequence #3000.
XX
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200062736-A2.
XX
PD 26-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US08020.
XX
PR 06-APR-1999; 99US-0127958.
XX
PA (UYEC-) UNIV EAST CAROLINA.
PA (NYCE/) NYCE J W.
XX
PI Nyce JW;
XX
DR WPI; 2000-679539/66.
XX
PT Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
PS Disclosure; Page 1416-1418; 1592pp; English.
XX
CC The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.
XX
SQ Sequence 6546 BP; 1603 A; 1691 C; 1500 G; 1751 T; 1 other;

Alignment Scores:

Pred. No.:	1.81e-187	Length:	6546
Score:	2142.00	Matches:	405
Percent Similarity:	99.26%	Conservative:	0
Best Local Similarity:	99.26%	Mismatches:	3
Query Match:	99.30%	Indels:	0
DB:	21	Gaps:	0

US-09-208-629F-6 (1-408) x AAF21433 (1-6546)

```
Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
              |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
```

```
Db        88 TGCTCCATGATTTTACAGATTTCATAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 147

Qy        21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       148 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 207

Qy        41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       208 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 267

Qy        61 GlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThrGly 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       268 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 327

Qy        81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       328 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 387

Qy       101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       388 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 447

Qy       121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       448 CTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTCTTC 507

Qy       141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       508 AGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTTCTT 567

Qy       161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       568 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTATTT 627

Qy       181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       628 GGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 687

Qy       201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       688 CTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTAC 747

Qy       221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       748 CGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACAGTT 807

Qy       241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       808 TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 867

Qy       261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       868 ATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTCTAT 927

Qy       281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       928 TACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTACTGC 987

Qy       301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       988 TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 1047

Qy       321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1048 GCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATTCTT 1107

Qy       341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1108 ATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATATCTC 1167

Qy       361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1168 ATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTCATG 1227

Qy       381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
             ||||||||||||||||||||||||||||||||||||||||||||      ||||||||||||||
Db      1228 TCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGAGAA 1287

Qy       401 GlnGlyGlnProSerGlnArgThr 408
             |||||||||||||||||||||||||
Db      1288 CAAGGACAGCCATCACAGAGAACG 1311
```

RESULT 6
AAV07375
ID AAV07375 standard; DNA; 1102 BP.
XX
AC AAV07375;
XX
DT 12-OCT-1998 (first entry)
XX
DE Human protease-activated receptor 3 (PAR3) genomic DNA.
XX
KW Protease-activated receptor 3; PAR3; thrombin receptor; human;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis; ds.
XX
OS Homo sapiens.
XX
PN WO9818456-A1.
XX
PD 07-MAY-1998.
XX
PF 29-OCT-1997; 97WO-US19732.
XX
PR 30-OCT-1996; 96US-0742440.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Connolly A, Coughlin SR, Ishihara H;
XX
DR WPI; 1998-271905/24.
XX
PT DNA encoding protease-activated receptor 3 - for detection of
PT specific agonists and antagonists, potentially useful for treating
PT e.g. thrombosis, atherosclerosis, inflammation etc.
XX
PS Claim 3; Page 41; 74pp; English.
XX
CC This sequence comprises genomic DNA (excluding exon 2) for human
CC protease-activated receptor 3 (PAR3), a cell surface protein which
CC is specifically activated by thrombin or a thrombin agonist,
CC thereby activating signalling events such as phosphoinositide
CC hydrolysis, calcium ion efflux and platelet aggregation. A cDNA
CC library was established from rat platelets and amplified (see
CC AAV07376-77). A product was identified that was expected to encode a
CC new G-protein coupled receptor, related to PAR1 or 2. This was
CC used to isolate human and murine DNA by a combination of PCR and
CC hybridisation techniques. Murine cDNA (see AAV07372), murine
CC genomic DNA (see AAV07373), human cDNA (see AAV07374) and human genomic
CC DNA sequences are provided, as well as the murine (see AAW51405) and
CC human (see AAW51406) PAR3 proteins. Also claimed are vectors, host
CC cells and an assay device. Host cells are used to screen compounds
CC for their ability to act as agonists or antagonists of the effects
CC of thrombin-PAR3 interaction. Agonists are used to treat wounds,
CC thrombosis, atherosclerosis, restenosis, inflammation and other
CC thrombin activated disorders. Antagonists (see AAW51415-21) are used
CC to control blood coagulation and thereby to treat heart attack and
CC stroke. They also mediate inflammatory and proliferative responses
CC to injury as occur in wound healing, atherosclerosis, restenosis,
CC pulmonary inflammation (ARDS) and glomerulosclerosis.
XX
SQ Sequence 1102 BP; 276 A; 286 C; 195 G; 345 T; 0 other;

Alignment Scores:

Pred. No.:	1.86e-170	Length:	1102
Score:	1947.00	Matches:	364
Percent Similarity:	99.46%	Conservative:	1
Best Local Similarity:	99.18%	Mismatches:	2
Query Match:	90.26%	Indels:	0
DB:	19	Gaps:	0

US-09-208-629F-6 (1-408) x AAV07375 (1-1102)

```
Qy         40 SerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACAGGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTT 60

Qy         60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACA 120

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTG 180

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAAATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATC 240

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACCTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTC 300

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCAGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTT 360

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTTTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTA 420

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTTGGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCC 480

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTCTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACC 540

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACA 600

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTTTTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCA 660

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTC 720

Qy        280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TATTACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTAC 780

Qy        300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrVal 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGCTATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTT 840

Qy        320 LysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIle 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGGCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATT 900

Qy        340 LeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyr 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATAT 960

Qy        360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTCATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTC 1020

Qy        380 MetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArg 399
              ||||||||||||||||||||||||||||||||||||||||||      ||||||||||||
Db       1021 ATGTCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGA 1080

Qy        400 GluGlnGlyGlnProSerGln 406
              |||||||||||||||||||||
Db       1081 GAACAAGGACAGCCATCACAG 1101
```

RESULT 7
AAV07372
ID AAV07372 standard; cDNA; 1224 BP.
XX
AC AAV07372;
XX
DT 12-OCT-1998 (first entry)
XX
DE Mouse protease-activated receptor 3 (PAR3) cDNA.
XX
KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 51..1160
FT /*tag= a
XX
PN WO9818456-A1.
XX
PD 07-MAY-1998.
XX
PF 29-OCT-1997; 97WO-US19732.
XX
PR 30-OCT-1996; 96US-0742440.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Connolly A, Coughlin SR, Ishihara H;
XX
DR WPI; 1998-271905/24.
DR P-PSDB; AAW51405.
XX
PT DNA encoding protease-activated receptor 3 - for detection of
PT specific agonists and antagonists, potentially useful for treating
PT e.g. thrombosis, atherosclerosis, inflammation etc.
XX
PS Claim 3; Page 38-39; 74pp; English.
XX
CC This cDNA clone codes for mouse protease-activated receptor 3
CC (PAR3) (see AAW51405), a cell surface protein which is specifically
CC activated by thrombin or a thrombin agonist, thereby activating
CC signalling events such as phosphoinositide hydrolysis, calcium ion
CC efflux and platelet aggregation. A cDNA library was establised
CC from rat platelets and amplified (primers given in AAV07376-77). A
CC product was identified that was expected to encode a new G-protein
CC coupled receptor, related to PAR1 or 2. This was used to isolate
CC human and murine DNA by a combination of PCR and hybridisation
CC techniques. Murine cDNA, murine genomic DNA (see AAV07373), human
CC cDNA (see AAV07374) and human genomic DNA (see AAV07375) sequences are
CC provided, as well as the murine and human (see AAW51406) PAR3
CC proteins. Also claimed are vectors, host cells and an assay
CC device. Host cells are used to screen compounds for their ability
CC to act as agonists or antagonists of the effects of thrombin-PAR3
CC interaction. Agonists are used to treat wounds, thrombosis,
CC atherosclerosis, restenosis, inflammation and other thrombin
CC activated disorders. Antagonists (see AAW51415-21) are used to

```
CC   control blood coagulation and thereby to treat heart attack and
CC   stroke.  They also mediate inflammatory and proliferative responses
CC   to injury as occur in wound healing, atherosclerosis, restenosis,
CC   pulmonary inflammation (ARDS) and glomerulosclerosis.
XX
SQ   Sequence 1224 BP; 299 A; 341 C; 239 G; 345 T; 0 other;

Alignment Scores:
Pred. No.:              6.87e-124      Length:         1224
Score:                  1443.50        Matches:        264
Percent Similarity:     83.74%         Conservative:   45
Best Local Similarity:  71.54%         Mismatches:     59
Query Match:            66.92%         Indels:         1
DB:                     19             Gaps:           1

US-09-208-629F-6 (1-408) x AAV07372 (1-1224)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||:::|||   |||||||||||   ||||||      ||||||||||||||||   |||
Db         33 GGGTCTCAGGACATCAAGATGAAAATCCTTATCTTGGTTGCAGCTGGGCTGCTGTTTCTG 92

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||   ||||||||||||:::   |||   ::::::|||   ||||||||||||
Db         93 CCAGTCACTGTTTGCCAAAGTGGCATA---AATGTTTCAGACAACTCAGCAAAGCCAACC 149

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluProProPheSer 73
              |||   ||||||:::|||   |||   |||   |||:::||||||||||||||   |||
Db        150 TTAACTATTAAGAGTTTTAATGGGGGTCCCCAAAATACCTTTGAAGAATTCCCACTTTCT 209

Qy         74 AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93
              :::||||||||||||||||||||   |||:::|||   :::|||||||||:::|||
Db        210 GACATAGAGGGCTGGACAGGAGCCACCACAACTATAAAAGCGGAGTGTCCCGAGGACAGT 269

Qy         94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThr 113
              |||   |||||||||   |||||||||:::|||||||||   ||||||||||||||||
Db        270 ATTTCAACTCTCCACGTGAATAATGCTACCATAGGATACCTGAGAAGTTCCTTAAGTACC 329

Qy        114 LysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaVal 133
              ::::::|||||||||||||||:::|||:::|||||||||||||||||||:::|||   |||
Db        330 CAAGTGATACCTGCCATCTATATCCTGCTGTTTGTGGTTGGTGTACCATCCAACATCGTG 389

Qy        134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153
              |||||||||   |||      ||||||:::||||||          :::|||:::||||||
Db        390 ACCCTGTGGAAACTCTCCTTAAGGACCAAATCCATCAGTCTGGTCATCTTTCACACCAAC 449

Qy        154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        450 CTGGCCATCGCAGATCTCCTTTTCTGTGTCACACTGCCATTTAAGATCGCCTACCATCTC 509

Qy        174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
              |||||||||||||||||||||||||||:::||||||   ||||||||||:::||||||
Db        510 AATGGCAACAACTGGGTATTTGGCGAGGTCATGTGCCGGATCACCACGGTCGTTTTCTAC 569

Qy        194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
              ||||||||||||||:::||||||:::|||   |||:::   ||||||||||||||||||
Db        570 GGCAACATGTACTGCGCTATCCTGATCCTCACTTGCATGGGCATCAACCGCTACCTGGCC 629

Qy        214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233
                  ||||||||||||||:::   |||||||||   ::::::::::|||:::   |||
Db        630 ACGGCTCACCCTTTCACATACCAGAAGCTGCCCAAACGCAGCTTCTCCTTGCTCATGTGT 689

Qy        234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
              |||:::||||||      ||||||||||||||||||||||   ||||||||||||||||
Db        690 GGCATAGTGTGGGTCATGGTTTTCTTATACATGCTGCCCTTTGTCATCCTGAAGCAGGAG 749

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
              |||:::||||||      :::||||||||||||||||||||||   :::   |||||||||
Db        750 TACCACCTCGTCCACTCAGAGATCACCACCTGCCACGATGTCGTCGACGCGTGCGAGTCC 809

Qy        274 SerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIlePro 293
                 |||   |||:::   |||||||||:::|||||||||||||||||||||||||||||||
Db        810 CCATCATCCTTCCGATTCTACTACTTCGTCTCCTTAGCATTCTTTGGGTTCCTCATCCCG 869

Qy        294 PheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHis 313
              ||||||:::||||||:::||||||      :::|||      |||   :::   |||
Db        870 TTTGTGATCATCATCTTCTGTTACACGACTCTCATCCACAAACTTAAATCAAAGGATCGG 929

Qy        314 ArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPhe 333
                 ||||||   |||:::||||||   ||||||||||||||||||||||||||||||||
Db        930 ATATGGCTGGGCTACATCAAGGCCGTCCTCCTCATCCTTGTGATTTTCACAATTTGCTTT 989

Qy        334 AlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAsp 353
              ||||||:::||||||||||||:::||||||||||||||||||||||:::||||||||||
Db        990 GCCCCCACCAACATCATACTCGTAATCCACCATGCCAACTACTACTACCACAATACCGAC 1049

Qy        354 GlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
                 |||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db       1050 AGCTTGTACTTTATGTATCTTATTGCTCTGTGCCTGGGGAGCCTGAATAGCTGCCTAGAT 1109

Qy        374 ProPheLeuTyrPheLeuMetSerLys 382
                 |||||||||||||||:::|||||||||
Db       1110 CCATTCCTTTACTTTGTCATGTCGAAA 1136

RESULT 8
AAZ50774
ID   AAZ50774 standard; DNA; 2409 BP.
XX
AC   AAZ50774;
XX
DT   31-MAY-2000  (first entry)
XX
DE   Human protease activated receptor-3 DNA.
XX
KW   Human; PAR-3; protease activated receptor;
KW   antisense molecule; PAR antibody; cytostatic; therapeutic;
KW   metastatic tumour cell; placental implantation; invasive cell; ds.
XX
OS   Homo sapiens.
XX
PN   WO200008150-A1.
XX
PD   17-FEB-2000.
XX
PF   05-FEB-1999;   99WO-IL00079.
XX
PR   07-AUG-1998;   98IL-0125698.
XX
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
PI   Bar-Shavit R;
XX
DR   WPI; 2000-205706/18.
XX
PT   Treating metastatic tumor cells useful for treating disorders involving
PT   placenta implantation in a female comprises administration of an
PT   antisense molecule complementary to an RNA sequence of a protease
PT   activated receptor protein -
XX
PS   Example 3; Fig 10; 46pp; English.
XX
CC   The patent discloses a method to treat metastatic tumour cells using
CC   an antisense molecule comprising a polynucleotide complementary to an
CC   RNA sequence of a protease activated receptor (PAR) protein, or an
CC   antibody capable of binding to a PAR protein. The antisense molecules and
CC   antibodies of PAR protein are also used to treat disorders associated
CC   with implantation of placenta. The present sequence is a human PAR-3
CC   DNA used for producing antisense molecules for treating invasive
CC   cells.
XX
SQ   Sequence 2409 BP; 598 A; 652 C; 454 G; 703 T; 2 other;

Alignment Scores:
Pred. No.:              1.65e-123      Length:         2409
Score:                  1443.50        Matches:        264
Percent Similarity:     83.74%         Conservative:   45
Best Local Similarity:  71.54%         Mismatches:     59
```

```
Query Match:            66.92%      Indels:        1
DB:                     21          Gaps:          1

US-09-208-629F-6 (1-408) x AAZ50774 (1-2409)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
Db        135 GGGTCTCAGGACATCAAGATGAAAATCCTTATCTTGGTTGCAGCTGGGCTGCTGTTTCTG 194

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
Db        195 CCAGTCACTGTTTGCCAAAGTGGCATA---AATGTTTCAGACAACTCAGCAAAGCCAACC 251

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluProProPheSer 73
Db        252 TTAACTATTAAGAGTTTTAATGGGGGTCCCCAAAATACCTTTGAAGAATTCCCACTTTCT 311

Qy         74 AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93
Db        312 GACATAGAGGGCTGGACAGGAGCCACCACAACTATAAAAGCGGAGTGTCCCGAGGACAGT 371

Qy         94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThr 113
Db        372 ATTTCAACTCTCCACGTGAATAATGCTACCATAGGATACCTGAGAAGTTCCTTAAGTACC 431

Qy        114 LysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaVal 133
Db        432 CAAGTGATACCTGCCATCTATATCCTGCTGTTTGTGGTTGGTGTACCATCCAACATCGTG 491

Qy        134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153
Db        492 ACCCTGTGGAAACTCTCCTTAAGGACCAAATCCATCAGTCTGGTCATCTTTCACACCAAC 551

Qy        154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
Db        552 CTGGCCATCGCAGATCTCCTTTTCTGTGTCACACTGCCATTTAAGATCGCCTACCATCTC 611

Qy        174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
Db        612 AATGGCAACAACTGGGTATTTGGCGAGGTCATGTGCCGGATCACCACGGTCGTTTTCTAC 671

Qy        194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
Db        672 GGCAACATGTACTGCGCTATCCTGATCCTCACTTGCATGGGCATCAACCGCTACCTGGCC 731

Qy        214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233
Db        732 ACGGCTCACCCTTTCACATACCAGAAGCTGCCCAAACGCAGCTTCTCCTTGCTCATGTGT 791

Qy        234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
Db        792 GGCATAGTGTGGGTCATGGTTTTCTTATACATGCTGCCCTTTGTCATCCTGAAGCAGGAG 851

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
Db        852 TACCACCTCGTCCACTCAGAGATCACCACCTGCCACGATGTCGTCGACGCGTGCGAGTCC 911

Qy        274 SerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIlePro 293
Db        912 CCATCATCCTTCCGATTCTACTACTTCGTCTCCTTAGCATTCTTTGGGTTCCTCATCCCG 971

Qy        294 PheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHis 313
Db        972 TTTGTGATCATCATCTTCTGTTACACGACTCTCATCCACAAACTTAAATCAAAGGATCGG 1031

Qy        314 ArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPhe 333
Db       1032 ATATGGCTGGGCTACATCAAGGCCGTCCTCCTCATCCTTGTGATTTTCACAATTTGCTTT 1091

Qy        334 AlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAsp 353
Db       1092 GCCCCCACCAACATCATACTCGTAATCCACCATGCCAACTACTACTACCACAATACCGAC 1151

Qy        354 GlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
Db       1152 AGCTTGTACTTTATGTATCTTATTGCTCTGTGCCTGGGGAGCCTGAATAGCTGCCTAGAT 1211

Qy        374 ProPheLeuTyrPheLeuMetSerLys 382
Db       1212 CCATTCCTTTACTTTGTCATGTCGAAA 1238
```

RESULT 9
AAV07373
ID AAV07373 standard; DNA; 1124 BP.
XX
AC AAV07373;
XX
DT 12-OCT-1998 (first entry)
XX
DE Mouse protease-activated receptor 3 (PAR3) genomic DNA.
XX
KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis; ds.
XX
OS Mus sp.
XX
PN WO9818456-A1.
XX
PD 07-MAY-1998.
XX
PF 29-OCT-1997; 97WO-US19732.
XX
PR 30-OCT-1996; 96US-0742440.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Connolly A, Coughlin SR, Ishihara H;
XX
DR WPI; 1998-271905/24.
XX
PT DNA encoding protease-activated receptor 3 - for detection of
PT specific agonists and antagonists, potentially useful for treating
PT e.g. thrombosis, atherosclerosis, inflammation etc.
XX
PS Claim 3; Page 39; 74pp; English.
XX
CC This sequence comprises genomic DNA for mouse protease-activated
CC receptor 3 (PAR3), a cell surface protein which is specifically
CC activated by thrombin or a thrombin agonist, thereby activating
CC signalling events such as phosphoinositide hydrolysis, calcium ion
CC efflux and platelet aggregation. A cDNA library was establised
CC from rat platelets and amplified (primers given in AAV07376-77). A
CC product was identified that was expected to encode a new G-protein
CC coupled receptor, related to PAR1 or 2. This was used to isolate
CC human and murine DNA by a combination of PCR and hybridisation
CC techniques. Murine cDNA (see AAV07372), murine genomic DNA, human
CC cDNA (see AAV07374) and human genomic DNA (see AAV07375) sequences are
CC provided, as well as the murine (see AAW51405) and human (see AAW51406)
CC PAR3 proteins. Also claimed are vectors, host cells and an assay
CC device. Host cells are used to screen compounds for their ability
CC to act as agonists or antagonists of the effects of thrombin-PAR3
CC interaction. Agonists are used to treat wounds, thrombosis,
CC atherosclerosis, restenosis, inflammation and other thrombin
CC activated disorders. Antagonists (see AAW51415-21) are used to
CC control blood coagulation and thereby to treat heart attack and
CC stroke. They also mediate inflammatory and proliferative responses
CC to injury as occur in wound healing, atherosclerosis, restenosis,
CC pulmonary inflammation (ARDS) and glomerulosclerosis.
XX
SQ Sequence 1124 BP; 251 A; 314 C; 210 G; 326 T; 23 other;

Alignment Scores:
Pred. No.: 3.48e-79 Length: 1124

```
Score:                  959.00        Matches:       185
Percent Similarity:     73.99%        Conservative:   34
Best Local Similarity:  62.50%        Mismatches:     53
Query Match:            44.46%        Indels:         25
DB:                     19            Gaps:           2

US-09-208-629F-6 (1-408) x AAV07373 (1-1124)

Qy         44 AsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaPro 63
Db         36 AATGTTTCAGACAACTCAGCAAAGCCAACCTTAACTATTAAGAGTTTTAATGGGGGTCCC 95

Qy         64 ProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIle 83
Db         96 CAAAATACCTTTGAAGAATTCNN---------------------------------- 118

Qy         84 ThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThr 103
Db        119 --------------------------NNNNNTACAACTCTCCATGTGAATAATGCTACC 151

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
Db        152 ATGGGATACCTGAGAAGTTCCTTAAGTACCAAAGTGATACCTGCCATCTACATCCTGGTG 211

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArg 143
Db        212 TTTGTGATTGGTGTACCAGCGAACATCGTGACCCTGTGGAAACTCTCCTCAAGGACCAAA 271

Qy        144 SerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        272 TCCATCTGTCTGGTCATCTTTCACACCAACCTGGCCATCGCGGATCTCCTTTTCTGTGTC 331

Qy        164 ThrLeuProPheLysIle-AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVa 183
Db        332 ACGCTGCCGTTTAAGATCNNCCTACCATCTCAATGGCAACAACTGGGTATTTGGCGAGGT 391

Qy        183 lLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer---IleLeuLe 202
Db        392 CATGTGCCGGATCACCACGGTCGTTTTCTACGGCAACATGTACTGCGCTANNNTCCTGAT 451

Qy        202 uLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGl 222
Db        452 CCTCACCTGCATGGGCATCAACCGCTACCTGGCCACGGCTCACCCTTTCACATACCAGAA 511

Qy        222 yLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLe 242
Db        512 GCTGCCCAAACGCAGCTTCTCCATGCTCATGTGTGGCATGGTGTGGGTCATGGTTTTCTT 571

Qy        242 uTyrMetLeuProPhePheIle-LeuLysGlnGluTyrTyrLeuValGlnProAspIleT 262
Db        572 ATACATGCTGCCCTTTGTCATCCNNNAAGCAGGAGTACCACCTCGTCCACTCCGAGATCA 631

Qy        262 hrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrP 282
Db        632 CCACCTGCCACGATGTCGTCGACGCGTGCGANTCCCCATCATCCTTCCGATTCTACTACT 691

Qy        282 heIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrA 302
Db        692 TCGTCTCCTTAGCATTCTTTGGGTTCCTCATCCCGTTTGTGATCATCATCTTCTGTTACA 751

Qy        302 laAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaS 322
Db        752 CGACTCTCATCCACAAACTTAAATCAAAAGATCNGATATGGCTGGGCTACATCAAGGCCG 811

Qy        322 erLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
Db        812 TCCTCCTCATCCTTGTGAATTTCACCATCTGCTTCCCCCCCACC 855

RESULT 10
AAQ84559
ID   AAQ84559 standard; cDNA; 2732 BP.
XX
AC   AAQ84559;
XX
DT   22-AUG-1995  (first entry)
XX
DE   Murine C140 receptor cDNA.
XX
KW   G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             73..1272
FT                   /*tag=  a
XX
PN   WO9503318-A.
XX
PD   02-FEB-1995.
XX
PF   26-JUL-1994;    94WO-US08536.
XX
PR   26-JUL-1993;    93US-0097938.
XX
PA   (CORT-) COR THERAPEUTICS.
XX
PI   Scarborough RM,  Sundelin J;
XX
DR   WPI; 1995-075182/10.
DR   P-PSDB; AAR66922.
XX
PT   New DNA encoding recombinant C140 receptor - and novel agonists
PT   and antagonists and specific antibodies with therapeutic and
PT   diagnostic applications.
XX
PS   Claim 1; Fig 10; 57pp; English.
XX
CC   A cDNA library from a mouse stomach was constructed in lambda gt10
CC   and screened with a probe encompassing the C140 genomic clone (see
CC   AAQ84557). A single phage clone was isolated and cut with EcoRI. The
CC   insert was cloned in pBluescript and pSG5 and sequenced. The
CC   complete nt sequence and deduced AA sequence is given in AAQ84559 &
CC   AAR66992. 5' RACE resulted in the addition of only 27 bps to the 5'
CC   end. The 5' end of the apparent coding region differs from the 5'
CC   end of the ORF of genomic DNA; it is believed that the 5' end of
CC   the cDNA sequence is correct.
XX
SQ   Sequence 2732 BP; 650 A; 676 C; 633 G; 773 T; 0 other;

Alignment Scores:
Pred. No.:              3.98e-46      Length:        2732
Score:                  606.50        Matches:       137
Percent Similarity:     53.23%        Conservative:  69
Best Local Similarity:  35.40%        Mismatches:    126
Query Match:            28.12%        Indels:        55
DB:                     16            Gaps:          12

US-09-208-629F-6 (1-408) x AAQ84559 (1-2732)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProPheSerAlaLeuGluGlyTrpThr 79
Db        205 GAAACCCAGCCTCCA-------------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97
Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
```

```
Db        280 ----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137

Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156

Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176

Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196

Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216

Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235

Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253

Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273

Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG---------------------- 765

Qy        274 SerSerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhe 287

Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307

Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla-----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321

Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341

Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361

Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381

Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387

Db       1123 AAAGATTTCAGGGATCACGCC 1143
```

RESULT 11
AAT32038
ID AAT32038 standard; cDNA; 2732 BP.
XX
AC AAT32038;
XX
DT 02-APR-1997 (first entry)
XX
DE Murine C140 receptor cDNA clone.
XX
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure; ss.
XX
OS Mus sp.
XX
FH Key   Location/Qualifiers
FT CDS   73..1272
FT   /*tag= a
FT sig_peptide   73..165
FT   /*tag= b
FT   /note= "the signal sequence differs from a
FT   signal sequence shown in a genomic DNA
FT   segment encoding this receptor (see AAT32036),
FT   the signal sequence given for this clone is
FT   believed to be the correct sequence"
FT mat_peptide   166..1269
FT   /*tag= c
XX
PN WO9623225-A1.
XX
PD 01-AUG-1996.
XX
PF 25-JAN-1996; 96WO-US01179.
XX
PR 25-JAN-1995; 95US-0390301.
XX
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Scarborough RM, Sundelin J;
XX
DR WPI; 1996-362813/36.
DR P-PSDB; AAW01954.
XX
PT Vector for expression C140 cell surface receptor in host cell -
PT useful to identify C140 agonist and antagonists, which are
PT antihypertensives and elevators of blood pressure, respectively
XX
PS Example 4; Fig 10A-B; 60pp; English.
XX
CC AAT32038 encodes the murine C140 receptor (C140R). The sequence may be
CC engineered so as to allow the recombinant expression of C140R in a
CC suitable host cell, i.e. by removing the native expression-control
CC sequence and replacing them with control sequences operable in the
CC host. Such a recombinant receptor can be expressed on the surface of
CC oocytes, this provides a good assay system for identifying
CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
CC receptor and a member of the "seven-pass" transmembrane receptor
CC superfamily (peptide chain of the receptor passes through the cell
CC membrane seven times, producing seven transmembrane regions within
CC the receptor molecule). The C140 receptor is involved in controlling
CC blood pressure. C140 antagonists (see AAW01942-W01951) are useful to
CC inhibit signalling from this receptor, resulting in an increase in
CC blood pressure and are therefore useful in pharmaceuticals for the
CC treatment of hypotension (low blood pressure). Conversely agonists
CC (see AAW01914-W01941) of C140 are useful in pharmaceuticals for the
CC treatment of hypertension (high blood pressure).
XX
SQ Sequence 2732 BP; 650 A; 676 C; 633 G; 773 T; 0 other;

Alignment Scores:

Pred. No.:	3.98e-46	Length:	2732
Score:	606.50	Matches:	137
Percent Similarity:	53.23%	Conservative:	69
Best Local Similarity:	35.40%	Mismatches:	126
Query Match:	28.12%	Indels:	55
DB:	17	Gaps:	12

US-09-208-629F-6 (1-408) x AAT32038 (1-2732)

```
Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41

Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144
```

```
Qy        42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db       145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy        60 ArgGlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThr 79
Db       205 GAAACCCAGCCTCCA-----------------------------------ATCACT 225

Qy        80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97
Db       226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy        98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
Db       280 ----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy       118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137
Db       316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy       138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
Db       376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy       157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db       436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy       177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db       496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy       197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
Db       556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy       217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
Db       616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy       236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
Db       664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy       254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
Db       724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG----------------- 765

Qy       274 SerSerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhe 287
Db       766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy       288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
Db       823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy       308 LeuAsnAla-----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321
Db       883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy       322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
Db       943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy       342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
Db      1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy       362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db      1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy       382 LysThr---ArgAsnHisSer 387
```

```
Db      1123 AAAGATTTCAGGGATCACGCC 1143
```

RESULT 12
AAQ84557
ID AAQ84557 standard; DNA; 1477 BP.
XX
AC AAQ84557;
XX
DT 22-AUG-1995 (first entry)
XX
DE Murine C140 receptor genomic DNA.
XX
KW G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 232..1188
FT /*tag= a
XX
PN WO9503318-A.
XX
PD 02-FEB-1995.
XX
PF 26-JUL-1994; 94WO-US08536.
XX
PR 26-JUL-1993; 93US-0097938.
XX
PA (CORT-) COR THERAPEUTICS.
XX
PI Scarborough RM, Sundelin J;
XX
DR WPI; 1995-075182/10.
DR P-PSDB; AAR66920.
XX
PT New DNA encoding recombinant C140 receptor - and novel agonists
PT and antagonists and specific antibodies with therapeutic and
PT diagnostic applications.
XX
PS Disclosure; Fig 1; 57pp; English.
XX
CC A mouse cosmid genomic library (obtd. from Dr R.A Wetsel, Washington
CC Univ. School of Medicine, St Louis, Missouri) was screened with two
CC 32P-labeled oligos corresp. to bp 190-249 and 742-801 of the bovine
CC substance K receptor cDNA. In one of the clones isolated (C140) the
CC hybridising region was localised to a 3.7 kb PstI fragment. This
CC fragment was subcloned into pBluescript vector. The hybridising and
CC adjacent regions were sequenced. The nt sequence and the deduced AA
CC sequence are given in AAQ84557 & AAR69920 respectively. The published
CC DNA sequence appears to have been printed with the last two lines
CC in the wrong order. This has been corrected in this GeneSeq entry.
XX
SQ Sequence 1477 BP; 341 A; 414 C; 320 G; 402 T; 0 other;

Alignment Scores:

Pred. No.:	6.44e-46	Length:	1477
Score:	600.50	Matches:	139
Percent Similarity:	52.25%	Conservative:	70
Best Local Similarity:	34.75%	Mismatches:	118
Query Match:	27.84%	Indels:	74
DB:	16	Gaps:	13

US-09-208-629F-6 (1-408) x AAQ84557 (1-1477)

```
Qy        18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuLeuProThr--- 36
Db       221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy        37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla------------ 50
Db       281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340
```

```
Qy          51 ----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
Db         341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy          65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db         400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy          85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db         418 ----------------------TCTGCGTCC------------------------- 426

Qy         105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db         427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy         125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db         484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy         145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db         544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy         164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db         604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy         184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db         664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy         204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db         724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy         224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
Db         775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy         241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db         832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy         261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db         892 ATCACCACCTGTCACGATGTG---------------------CTGCCTGAGGAGGTATTG 930

Qy         281 -----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db         931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy         295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
Db         991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy         311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db        1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy         329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db        1111 TACTTCATCTGCTTTCGTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy         349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db        1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy         369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db        1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290
```

RESULT 13

```
AAT32036
ID   AAT32036 standard; DNA; 1477 BP.
XX
AC   AAT32036;
XX
DT   01-APR-1997  (first entry)
XX
DE   Murine C140 receptor-encoding DNA.
XX
KW   C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW   antagonist; hypertension; hypotension; blood pressure; ds.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             232..1419
FT                   /*tag=  a
FT   sig_peptide     232..312
FT                   /*tag=  b
FT                   /note= "putative signal sequence, differs from
FT                          signal sequence shown in a cDNA clone of
FT                          this receptor (see AAT32038), the signal
FT                          sequence given for the cDNA clone is
FT                          believed to be the correct sequence"
FT   mat_peptide     313..1416
FT                   /*tag=  c
XX
PN   WO9623225-A1.
XX
PD   01-AUG-1996.
XX
PF   25-JAN-1996;   96WO-US01179.
XX
PR   25-JAN-1995;   95US-0390301.
XX
PA   (CORT-) COR THERAPEUTICS INC.
XX
PI   Scarborough RM,  Sundelin J;
XX
DR   WPI; 1996-362813/36.
DR   P-PSDB; AAW01952.
XX
PT   Vector for expression C140 cell surface receptor in host cell  -
PT   useful to identify C140 agonist and antagonists, which are
PT   antihypertensives and elevators of blood pressure, respectively
XX
PS   Example 1; Fig 1A-B; 60pp; English.
XX
CC   AAT32036 encodes the murine C140 receptor (C140R), including a tentative
CC   signal sequence. The sequence may be engineered so as to allow the
CC   recombinant expression of C140R in a suitable host cell, i.e. by
CC   removing the native expression-control sequence and replacing them with
CC   control sequences operable in the host. Such a recombinant receptor
CC   can be expressed on the surface of oocytes, this provides a good assay
CC   system for identifying agonists/antagonists of C140R. The C140
CC   receptor is a G-protein linked receptor and a member of the "seven-
CC   pass" transmembrane receptor superfamily (peptide chain of the
CC   receptor passes through the cell membrane seven times, producing
CC   seven transmembrane regions within the receptor molecule). The C140
CC   receptor is involved in controlling blood pressure. C140 antagonists
CC   (see AAW01942-W01951) are useful to inhibit signalling from this
CC   receptor, resulting in an increase in blood pressure and are therefore
CC   useful in pharmaceuticals for the treatment of hypotension (low blood
CC   pressure). Conversely agonists (see AAW01914-W01941) of C140 are useful
CC   in pharmaceuticals for the treatment of hypertension (high blood
CC   pressure).
XX
SQ   Sequence 1477 BP; 341 A; 415 C; 319 G; 402 T; 0 other;

Alignment Scores:
Pred. No.:              6.44e-46        Length:         1477
Score:                  600.50          Matches:        139
Percent Similarity:     52.13%          Conservative:   69
```

```
Best Local Similarity:  34.84%        Mismatches:    120
Query Match:            27.84%        Indels:         72
DB:                     17            Gaps:           12

US-09-208-629F-6 (1-408) x AAT32036 (1-1477)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla---------- 50
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 ----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db        400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db        418 ----------------------TCTGCGTCC-------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 Pro------LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
Db        775 CCCAGGAAGAAGGCAAACATCGCCGTTGGCTCCTCCTTGGCAATCTGGCTCCTGATTTTT 834

Qy        242 LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db        835 CTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAACATC 894

Qy        262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr--- 280
Db        895 ACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTGGTG 933

Qy        281 --------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheVal 295
Db        934 GGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCCCTC 993

Qy        296 LeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla-------------- 310
Db        994 CTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGATGAA 1053

Qy        311 ---TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePhe 329
Db       1054 CACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATGTAC 1113

Qy        330 ThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyr 349
Db       1114 TTCATCTGCTTTCGTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAAACC 1173

Qy        350 AsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
Db       1174 CAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTCAAC 1233

Qy        370 SerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db       1234 AGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290

RESULT 14
AAC60319
ID   AAC60319 standard; DNA; 1451 BP.
XX
AC   AAC60319;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Human PAR-2 DNA.
XX
KW   PAR-2; protease activated receptor-2; ECL-2; inflammatory disease;
KW   asthma; chronic obstructive pulmonary; arthritis;
KW   inflammatory bowel; psoriasis; eczema; multiple sclerosis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200063371-A1.
XX
PD   26-OCT-2000.
XX
PF   17-APR-2000; 2000WO-GB01455.
XX
PR   15-APR-1999;   99GB-0008513.
XX
PA   (UYSO-) UNIV SOUTHAMPTON.
XX
PI   Walls AF,  Palmer K,  Compton SJ,  Cairns JA,  Gough AC;
XX
DR   WPI; 2000-679599/66.
XX
PT   Protease activated receptor 2 variants useful for treating inflammatory
PT   diseases such as asthma, arthritis and psoriasis, and as hypertensives,
PT   has reduced sensitivity to trypsin -
XX
PS   Disclosure; Page 54-55; 59pp; English.
XX
CC   The present invention relates to a variant protease activated
CC   receptor 2 (PAR-2). The invention is useful for identifying an
CC   individual having a polymorphism in the ECL-2 region of one or
CC   both PAR-2 gene alleles. The invention may be used to develop treatments
CC   for inflammatory diseases such as asthma, chronic obstructive
CC   pulmonary diseases, arthritis, inflammatory bowel diseases, psoriasis
CC   and eczema, multiple sclerosis and to raise blood pressure.
XX
SQ   Sequence 1451 BP; 310 A; 389 C; 346 G; 406 T; 0 other;

Alignment Scores:
Pred. No.:              1.24e-44        Length:         1451
Score:                  586.50          Matches:        141
Percent Similarity:     50.61%          Conservative:   67
Best Local Similarity:  34.31%          Mismatches:     146
Query Match:            27.19%          Indels:         57
DB:                     21              Gaps:           14

US-09-208-629F-6 (1-408) x AAC60319 (1-1451)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
Db        139 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 198
```

```
Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db        199 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 237

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
Db        238 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA--------------- 282

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
Db        283 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 324

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db        325 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 354

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
Db        355 ACTGGAAAACTGACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGT 414

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
Db        415 TTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 474

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db        475 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 534

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db        535 TTGAAGATTGCCTATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 594

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
Db        595 GTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 654

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
Db        655 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 711

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
Db        712 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCT 771

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db        772 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 831

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr------------------ 280
Db        832 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 870

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db        871 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 930

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla-----------------TyrAspHisArg 314
Db        931 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 990

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
Db        991 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACT 1050

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db       1051 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1101

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db       1102 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1161

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
```

```
Db       1162 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1221

Qy        390 TyrLeuThrLys******AsnAspLeuArgGlu 400
Db       1222 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1254

RESULT 15
AAQ84558
ID   AAQ84558 standard; DNA; 1255 BP.
XX
AC   AAQ84558;
XX
DT   22-AUG-1995  (first entry)
XX
DE   Human C140 receptor genomic DNA.
XX
KW   G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             56..1197
FT                   /*tag=  a
XX
PN   WO9503318-A.
XX
PD   02-FEB-1995.
XX
PF   26-JUL-1994;   94WO-US08536.
XX
PR   26-JUL-1993;   93US-0097938.
XX
PA   (CORT-) COR THERAPEUTICS.
XX
PI   Scarborough RM,  Sundelin J;
XX
DR   WPI; 1995-075182/10.
DR   P-PSDB; AAR66921.
XX
PT   New DNA encoding recombinant C140 receptor - and novel agonists
PT   and antagonists and specific antibodies with therapeutic and
PT   diagnostic applications.
XX
PS   Disclosure; Fig 2; 57pp; English.
XX
CC   The availability of genomic DNA encoding the mouse protease
CC   C140 receptor (see Q84557) permitted the retrieval of the corresp.
CC   human gene. A human genomic library cloned in the vector EMBL3 was
CC   screened using the entire coding region of the murine clone as a
CC   probe. The recovered human gene including the DNA sequence and the
CC   deduced AA sequence are shown in Q84558 & R66921. Subsequent
CC   experiments indicated that the human C140 gene is located in the
CC   same region of the long arm of chromosome number 5 (5q12-5q13)
CC   as has been reported for the human thrombin receptor gene.
XX
SQ   Sequence 1255 BP; 294 A; 320 C; 260 G; 381 T; 0 other;

Alignment Scores:
Pred. No.:              1.27e-44        Length:         1255
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     16              Gaps:           10

US-09-208-629F-6 (1-408) x AAQ84558 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
```

```
Db        260 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139

Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158

Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178

Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198

Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218

Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238

Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258

Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278

Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292

Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310

Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 ------------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326

Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346

Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363

Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383

Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400

Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 16
AAT32037
ID   AAT32037 standard; DNA; 1255 BP.
XX
AC   AAT32037;
XX
DT   01-APR-1997  (first entry)
XX
DE   Human C140 receptor-encoding DNA.
XX
KW   C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW   antagonist; hypertension; hypotension; blood pressure; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             56..1197
FT                   /*tag=  a
FT   sig_peptide     56..136
FT                   /*tag=  b
FT                   /note= "putative signal sequence, differs from
FT                          signal sequence shown in a cDNA clone of
FT                          this receptor (see AAT32039 - this clone also
FT                          encodes a mature receptor having four amino
FT                          acid changes, when compared to this sequence)"
FT   mat_peptide     137..1194
FT                   /*tag=  c
XX
PN   WO9623225-A1.
XX
PD   01-AUG-1996.
XX
PF   25-JAN-1996;   96WO-US01179.
XX
PR   25-JAN-1995;   95US-0390301.
XX
PA   (CORT-) COR THERAPEUTICS INC.
XX
PI   Scarborough RM,  Sundelin J;
XX
DR   WPI; 1996-362813/36.
DR   P-PSDB; AAW01953.
XX
PT   Vector for expression C140 cell surface receptor in host cell  -
PT   useful to identify C140 agonist and antagonists, which are
PT   antihypertensives and elevators of blood pressure, respectively
XX
PS   Example 2; Fig 2A-B; 60pp; English.
XX
CC   AAT32037 encodes the human C140 receptor (C140R), including a tentative
CC   signal sequence. The sequence may be engineered so as to allow the
CC   recombinant expression of C140R in a suitable host cell, i.e. by
CC   removing the native expression-control sequence and replacing them with
CC   control sequences operable in the host. Such a recombinant receptor
CC   can be expressed on the surface of oocytes, this provides a good assay
CC   system for identifying agonists/antagonists of C140R. The C140
CC   receptor is a G-protein linked receptor and a member of the "seven-
CC   pass" transmembrane receptor superfamily (peptide chain of the
CC   receptor passes through the cell membrane seven times, producing
CC   seven transmembrane regions within the receptor molecule). The C140
CC   receptor is involved in controlling blood pressure. C140 antagonists
CC   (see AAW01942-W01951) are useful to inhibit signalling from this
CC   receptor, resulting in an increase in blood pressure and are therefore
CC   useful in pharmaceuticals for the treatment of hypotension (low blood
CC   pressure). Conversely agonists (see AAW01914-W01941) of C140 are useful
CC   in pharmaceuticals for the treatment of hypertension (high blood
CC   pressure).
XX
SQ   Sequence 1255 BP; 294 A; 320 C; 260 G; 381 T; 0 other;

Alignment Scores:
Pred. No.:              1.27e-44        Length:         1255
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     17              Gaps:           10

US-09-208-629F-6 (1-408) x AAT32037 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99

Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT---------- 259
```

```
Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                                  ||||||      |||:::|||      :::|||   :::
Db        260 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   :::|||||||||||||||:::|||:::|||   :::   ||||||:::
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                |||   |||      |||||||||:::||||||
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
              |||      :::      |||   ||||||||||||||:::::|||||||||||
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              ::::::||||||   ||||||                ||||||||||||||||||||
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
              :::          |||          |||:::                 :::|||
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
                 :::   |||      :::|||   :::::::::||||||      ::::::
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
              :::||||||||||||||||||||                          |||   |||
Db        719 CTGAACATCACGACCTGTCATGATGTT----------          ----------TTGCCTGAGCAG 757

Qy        279 LeuTyr----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              |||                ||||||:::||||||||         ||||||
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
              |||      |||          |||      :::||||||   |||   :::
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                        :::   :::          :::|||      :::   :::|||
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
              ::::::   |||||||||   |||||||||::::::|||:::::|||
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
              |||:::          :::   |||          :::|||   :::|||::::::||||||
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   :::|||||||||||:::||||||||:::|||:::   :::|||
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                 |||:::|||:::      |||   |||   :::          :::::::::
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 17
AAZ50773
ID   AAZ50773 standard; DNA; 1289 BP.
XX
AC   AAZ50773;
XX
DT   31-MAY-2000  (first entry)
XX
DE   Human protease activated receptor-2 DNA.
XX
KW   Human; PAR-2; protease activated receptor;
KW   antisense molecule; PAR antibody; cytostatic; therapeutic;
KW   metastatic tumour cell; placental implantation; invasive cell; ds.
XX
OS   Homo sapiens.
XX
PN   WO200008150-A1.
XX
PD   17-FEB-2000.
XX
PF   05-FEB-1999;   99WO-IL00079.
XX
PR   07-AUG-1998;   98IL-0125698.
XX
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
PI   Bar-Shavit R;
XX
DR   WPI; 2000-205706/18.
XX
PT   Treating metastatic tumor cells useful for treating disorders involving
PT   placenta implantation in a female comprises administration of an
PT   antisense molecule complementary to an RNA sequence of a protease
PT   activated receptor protein -
XX
PS   Example 3; Fig 9; 46pp; English.
XX
CC   The patent discloses a method to treat metastatic tumour cells using
CC   an antisense molecule comprising a polynucleotide complementary to an
CC   RNA sequence of a protease activated receptor (PAR) protein, or an
CC   antibody capable of binding to a PAR protein. The antisense molecules and
CC   antibodies of PAR protein are also used to treat disorders associated
CC   with implantation of placenta. The present sequence is a human PAR-2
CC   DNA used for producing antisense molecules for treating invasive
CC   cells.
XX
SQ   Sequence 1289 BP; 300 A; 320 C; 270 G; 399 T; 0 other;

Alignment Scores:
Pred. No.:              1.31e-44        Length:         1289
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     21              Gaps:           10

US-09-208-629F-6 (1-408) x AAZ50773 (1-1289)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||   |||:::                     |||   |||||||||
Db        168 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 215

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                                  ||||||      |||:::|||      :::|||   :::
Db        216 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 257

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   :::|||||||||||||||:::|||:::|||   :::   ||||||:::
Db        258 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 317

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                |||   |||      |||||||||:::||||||
Db        318 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 377

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
              |||      :::      |||   ||||||||||||||:::::|||||||||||
Db        378 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 437

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              ::::::||||||   ||||||                ||||||||||||||||||||
Db        438 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 497
```

```
Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db        498 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 557

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
                 :::           |||         |||:::                 :::|||
Db        558 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 614

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
                 :::   |||      :::|||   :::::::::||||||      ::::::
Db        615 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 674

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
                 :::||||||||||||||||||||                          |||   |||
Db        675 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 713

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPheGlyPheLeuIle 292
              |||                  ||||||:::|||||||||          ||||||
Db        714 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 773

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
              |||      |||           |||      :::||||||   |||   :::
Db        774 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 833

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                          :::   :::            :::|||      :::   :::|||
Db        834 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 893

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
                 ::::::   |||||||||   |||||||||::::::|||:::::::|||
Db        894 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 944

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
              |||:::          :::   |||         :::|||   :::|||::::::||||||
Db        945 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1004

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   :::|||||||||||||:::||||||||:::|||:::   :::|||
Db       1005 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1064

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                 |||:::|||:::      |||   |||   :::           :::::::::
Db       1065 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1121

RESULT 18
AAQ84560
ID   AAQ84560 standard; cDNA; 1414 BP.
XX
AC   AAQ84560;
XX
DT   22-AUG-1995  (first entry)
XX
DE   Human  C140 receptor cDNA.
XX
KW   G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             50..1243
FT                   /*tag=  a
XX
PN   WO9503318-A.
XX
PD   02-FEB-1995.
XX
PF   26-JUL-1994;   94WO-US08536.
XX
PR   26-JUL-1993;   93US-0097938.
XX
PA   (CORT-) COR THERAPEUTICS.
XX
PI   Scarborough RM,  Sundelin J;
XX
DR   WPI; 1995-075182/10.
DR   P-PSDB; AAR66923.
XX
PT   New DNA encoding recombinant C140 receptor - and novel agonists
PT   and antagonists and specific antibodies with therapeutic and
PT   diagnostic applications.
XX
PS   Claim 1; Fig 11; 57pp; English.
XX
CC   A human intestinal tumour cDNA library was subjected to PCR using
CC   primers designed from the genomic clone (see AAQ84558) and the
CC   amplified fragment was cloned in pSG5 and sequenced. There are
CC   four AA differences between the cDNA encoded sequence and that
CC   encoded by the genomic DNA. The genomic DNA sequence and deduced
CC   AA sequence are given in AAQ84560 & AAR66923.
XX
SQ   Sequence 1414 BP; 335 A; 361 C; 309 G; 409 T; 0 other;

Alignment Scores:
Pred. No.:              2.35e-43        Length:         1414
Score:                  572.50          Matches:        140
Percent Similarity:     49.64%          Conservative:   66
Best Local Similarity:  33.73%          Mismatches:     152
Query Match:            26.54%          Indels:         57
DB:                     16              Gaps:           14

US-09-208-629F-6 (1-408) x AAQ84560 (1-1414)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
              |||   |||:::|||      :::           :::   |||   |||      |||
Db         41 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 100

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
              |||   |||                     |||   ||||||            ||||||
Db        101 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 139

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
                    :::|||          |||          |||
Db        140 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA--------------- 184

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
                              |||   :::   |||      |||   |||:::
Db        185 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                       |||   |||||||||                                 |||
Db        227 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 256

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
                       |||:::|||      :::|||   :::|||   :::||||||   ||||||
Db        257 GCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGT 316

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPheArgThrArgSerIleCysThr 147
              :::|||:::|||   :::   ||||||:::      |||||||||:::
Db        317 TTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 376

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||   |||      |||||||||:::||||||   |||      :::      |||
Db        377 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 436

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
                 ||||||||||||||:::::||||||||||:::::||||||   ||||||
Db        437 TTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 496

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                          ||||||   ||||||||||||||||||||   :::   |||:::
Db        497 GTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 556

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
              |||:::   ||||||      ||||||:::|||      :::                |||
Db        557 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 613
```

```
Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
Db        614 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCT 673

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db        674 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 733

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr---------------------- 280
Db        734 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db        773 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 832

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla--------------------TyrAspHisArg 314
Db        833 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
Db        893 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACT 952

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db        953 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1003

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db       1004 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1063

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
Db       1064 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1123

Qy        390 TyrLeuThrLys******AsnAspLeuArgGluGlnGlyGlnPro 404
Db       1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168

RESULT 19
AAT32039
ID   AAT32039 standard; cDNA; 1414 BP.
XX
AC   AAT32039;
XX
DT   02-APR-1997  (first entry)
XX
DE   Human C140 receptor cDNA clone.
XX
KW   C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW   antagonist; hypertension; hypotension; blood pressure; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             50..1243
FT                   /*tag=  a
FT   sig_peptide     50..130
FT                   /*tag=  b
FT                   /note= "the signal sequence, differs from the
FT                          signal sequence shown in a genomic clone of
FT                          this receptor (see AAT32037). The cDNA clone
FT                          also encodes a mature receptor having four
FT                          amino acid changes, when compared to the
FT                          genomic clone"
FT   mat_peptide     131..1240
FT                   /*tag=  c
XX
PN   WO9623225-A1.
XX
PD   01-AUG-1996.
XX
PF   25-JAN-1996;   96WO-US01179.
XX
PR   25-JAN-1995;   95US-0390301.
XX
PA   (CORT-) COR THERAPEUTICS INC.
XX
PI   Scarborough RM,  Sundelin J;
XX
DR   WPI; 1996-362813/36.
DR   P-PSDB; AAW01955.
XX
PT   Vector for expression C140 cell surface receptor in host cell  -
PT   useful to identify C140 agonist and antagonists, which are
PT   antihypertensives and elevators of blood pressure, respectively
XX
PS   Example 5; Fig 11A-B; 60pp; English.
XX
CC   AAT32039 encodes the human C140 receptor (C140R). The sequence may be
CC   engineered so as to allow the recombinant expression of C140R in a
CC   suitable host cell, i.e. by removing the native expression-control
CC   sequence and replacing them with control sequences operable in the
CC   host. Such a recombinant receptor can be expressed on the surface of
CC   oocytes, this provides a good assay system for identifying
CC   agonists/antagonists of C140R. The C140 receptor is a G-protein linked
CC   receptor and a member of the "seven-pass" transmembrane receptor
CC   superfamily (peptide chain of the receptor passes through the cell
CC   membrane seven times, producing seven transmembrane regions within the
CC   receptor molecule). The C140 receptor is involved in controlling blood
CC   pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
CC   signalling from this receptor, resulting in an increase in blood
CC   pressure and are therefore useful in pharmaceuticals for the treatment
CC   of hypotension (low blood pressure). Conversely agonists (see AAW01914-
CC   AAW01941) of C140 are useful in pharmaceuticals for the treatment of
CC   hypertension (high blood pressure).
XX
SQ   Sequence 1414 BP; 335 A; 361 C; 309 G; 409 T; 0 other;

Alignment Scores:
Pred. No.:              2.35e-43        Length:         1414
Score:                  572.50          Matches:        140
Percent Similarity:     49.64%          Conservative:   66
Best Local Similarity:  33.73%          Mismatches:     152
Query Match:            26.54%          Indels:         57
DB:                     17              Gaps:           14

US-09-208-629F-6 (1-408) x AAT32039 (1-1414)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
Db         41 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 100

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db        101 GCCTCTCTC----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 139

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
Db        140 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA-------------- 184

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
Db        185 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db        227 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 256

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
Db        257 GCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGT 316

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
```

```
Db        317 TTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 376

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166

Db        377 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 436

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186

Db        437 TTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 496

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206

Db        497 GTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 556

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226

Db        557 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 613

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246

Db        614 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCT 673

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266

Db        674 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 733

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr------------------ 280

Db        734 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300

Db        773 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 832

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla------------------TyrAspHisArg 314

Db        833 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334

Db        893 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACT 952

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354

Db        953 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1003

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371

Db       1004 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1063

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389

Db       1064 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1123

Qy        390 TyrLeuThrLys******AsnAspLeuArgGluGlnGlyGlnPro 404

Db       1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168

RESULT 20
AAX90979
ID   AAX90979 standard; cDNA; 1361 BP.
XX
AC   AAX90979;
XX
DT   25-JAN-2000  (first entry)
XX
DE   Mouse protease-activated receptor 4 (PAR4) cDNA.
XX
KW   Mouse protease-activated receptor 4 cDNA; PAR4; thrombin receptor;
KW   phosphoinositide hydrolysis; calcium efflux; platelet aggregation;
KW   platelet activation; thrombin agonist; thrombin antagonist; therapeutic;
KW   wound; blood coagulation; heart attack; stroke; inflammatory response;
KW   proliferative response; atherosclerosis; restenosis; glomerulosclerosis;
KW   pulmonary inflammation; ARDS; adult respiratory distress syndrome; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1191
FT                   /*tag=  a
FT                   /product= "PAR4"
XX
PN   WO9943809-A2.
XX
PD   02-SEP-1999.
XX
PF   11-FEB-1999;   99WO-US02983.
XX
PR   27-FEB-1998;   98US-0032397.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Coughlin SR,  Kahn M;
XX
DR   WPI; 1999-619953/53.
DR   P-PSDB; AAY15081.
XX
PT   DNA molecules encoding protease-activated receptor 4, useful in
PT   compound assays for thrombin agonist and antagonist activity -
XX
PS   Claim 3; Fig 1; 69pp; English.
XX
CC   The present sequence is a cDNA encoding protease-activated receptor 4
CC   (PAR4) and derived from a mouse 14-15 days embryo library. PAR4 is a G
CC   protein-coupled thrombin receptor expressed on cell surface. It is
CC   activated by thrombin and mediates signalling events e.g.
CC   phosphoinositide hydrolysis, calcium efflux and platelet aggregation.
CC   The receptor is highly expressed in spleen cells and likely to be
CC   involved in thrombin-mediated activation of platelets and other
CC   haematopoietic cells. It is used for screening novel thrombin agonists
CC   or antagonists. The agonists are used as therapeutics to treat wounds,
CC   promote clotting and as reagents to activate platelets in diagnostic
CC   tests. Antagonists are used to control blood coagulation, treat heart
CC   attacks and strokes, and block inflammatory and proliferative responses
CC   that occur in normal wound healing and variety of diseases including
CC   atherosclerosis, restenosis, pulmonary inflammation (ARDS) and
CC   glomerulosclerosis.
XX
SQ   Sequence 1361 BP; 234 A; 416 C; 394 G; 317 T; 0 other;

Alignment Scores:
Pred. No.:              1.37e-39        Length:         1361
Score:                  531.50          Matches:        112
Percent Similarity:     53.82%          Conservative:   57
Best Local Similarity:  35.67%          Mismatches:     128
Query Match:            24.64%          Indels:         17
DB:                     20              Gaps:           2

US-09-208-629F-6 (1-408) x AAX90979 (1-1361)

Qy         84 ThrValLysIleLysCysProGluGluSer---------------------------- 93

Db        130 ACAGTAGAACTCAAGGAGCCGAAGTCCTCAGACAAGCCTAATCCACGAGGCTACCCGGGC 189

Qy         94 -----------------AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107

Db        190 AAATTCTGTGCCAACGACAGTGACACGCTGGAGCTCCCGGCCAGCTCTCAAGCACTGCTG 249

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127

Db        250 CTGGGGTGGGTATCCACAAGGCTGGTACCTGCCCTCTATGGGCTTGTGGTGGCTGTGGGG 309

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147

Db        310 CTGCCTGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACAAGGGTGCCACGCCTGCCATCC 369
```

```
Qy        148 ThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPhe 167
              |||:::     |||||||||||:::||||||   |||      :::      |||
Db        370 ACCATTCTGCTCACGAACCTGGCAGTGGCTGATTCGCTGTTGGCCCTGGTGCCGCCACCA 429

Qy        168 LysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAla 187
              ::::::|||||||||||   |||      |||   |||||||||      ||||||
Db        430 CGACTGGCTTACCACTTGCGTGGCCAGCGCTGGCCATTTGGTGAGGCTGCCTGCCGGGTG 489

Qy        188 ThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSer 207
              |||         ||||||:::||||||   |||:::||||||||||||   :::|||
Db        490 GCCACAGCTGCCCTCTATGGCCACATGTATGGTTCAGTGTTGCTGCTGGCTGCAGTCAGC 549

Qy        208 IleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThr 227
              ::::::|||||||||||:::|||||||||          |||   |||
Db        550 TTGGACAGATACCTGGCCCTGGTGCATCCTTTGCGGGCCCGTGCACTGCGTGGTCAACGC 609

Qy        228 TyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhe 247
                        |||   :::   |||   :::                |||||||
Db        610 CTCACTACTGGACTCTGTTTGGTGGCCTGGCTCTCTGCAGCCACCCTGGCCTTGCCTCTC 669

Qy        248 PheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspVal 267
              :::   :::|||   :::   |||         |||         |||||||||
Db        670 ACTCTGCATCGGCAGACCTTCCGATTAGCTGGCTCCGATCGCATGCTGTGTCATGATGCG 729

Qy        268 HisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhe 287
                        |||   :::|||   ::::::       ||||||   ||||||
Db        730 CTGCCCCTGACTGAGCAGACCTCCCACTGGAGA---CCGGCCTTCATCTGCCTGGCTGTC 786

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
              |||      :::|||   :::   :::      ||||||   |||   :::|||
Db        787 CTGGGCTGCTTCGTGCCACTGCTGGCCATGGGCCTGTGCTATGGAGCCACCCTTCGTGCA 846

Qy        308 LeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuVal 327
              |||   |||         |||:::         ::::::   :::   |||:::|||
Db        847 CTGGCGGCCAATGGCCAGCGCTACAGCCATGCACTCAGACTGACAGCCCTGGTACTGTTC 906

Qy        328 IlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyr 347
                        |||   |||||||||::::::|||::::::|||::::::|||
Db        907 TCGGCAGTGGCCTCTTTCACACCTAGCAATGTGCTGCTGGTGCTGCACTATTCAAACCCG 966

Qy        348 TyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySer 367
                              ||||||      |||:::   :::|||   |||   :::
Db        967 AGCCCTGAGGCCTGGGGCAATCTCTATGGAGCCTATGTGCCCAGCCTGGCACTCAGCACC 1026

Qy        368 LeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
              |||||||||||:::||||||||:::|||:::   :::|||
Db       1027 CTCAACAGCTGCGTAGACCCTTTCATCTACTACTATGTGTCC 1068
```

```
RESULT 21
AAZ32747
ID   AAZ32747 standard; cDNA; 4895 BP.
XX
AC   AAZ32747;
XX
DT   31-JAN-2000  (first entry)
XX
DE   Human protease-activated receptor PAR4 cDNA.
XX
KW   Protease-activated receptor; PAR4; ZCHEMR2; G protein coupled;
KW   cellular signalling; protease; tethered ligand; N-terminal;
KW   proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
KW   antagonist; cellular response; physiological response; clotting pathway;
KW   platelet; proliferation; differentiation; mediation;
KW   inflammatory process; vascular injury; chemotaxis; mitogenesis;
KW   growth factor; production; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             176..1333
FT                   /*tag=  a
FT                   /product= "Human protease-activated receptor PAR4"
FT   5'UTR           1..175
FT                   /*tag=  b
FT   sig_peptide     176..226
FT                   /*tag=  c
FT   mat_peptide     227..1330
FT                   /*tag=  d
FT                   /product= "Mature human PAR4"
FT   3'UTR           1333..4895
FT                   /*tag=  e
FT                   /note= "GC-rich"
XX
PN   WO9950415-A2.
XX
PD   07-OCT-1999.
XX
PF   31-MAR-1999;   99WO-US07100.
XX
PR   01-APR-1998;   98US-0053866.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Xu W,  Presnell SR,  Yee DP,  Foster DC;
XX
DR   WPI; 1999-633640/54.
DR   P-PSDB; AAY50135.
XX
PT   Novel protease activated receptor 4, useful for screening for
PT   (ant)agonists for promoting the proliferation and/or differentiation of
PT   platelets and in mediating inflammatory events  -
XX
PS   Claim 1; Page 76-80; 85pp; English.
XX
CC   This sequence represents a human protease-activated receptor PAR4 (also
CC   referred to as ZCHEMR2) cDNA which was identified from EST (expressed
CC   sequence tag) sequences with homology to the three known protease-
CC   activated receptors (PAR1, PAR2 and PAR3). Protease-activated receptors
CC   (PARs) are a subfamily of G protein coupled receptors which are capable
CC   of mediating cellular signalling in response to proteases (e.g.,
CC   thrombin). They are characterised by a tethered peptide ligand at the
CC   extracellular N-terminus that is generated by proteolysis. PAR4 is
CC   activated by thrombin or trypsin cleavage at Arg47/Gly48, which generates
CC   a new N-terminus corresponding to the tethered ligand (a hexapeptide).
CC   Agonists of PAR4 are useful for upregulating cellular or physiological
CC   responses whereas antagonists are used to downregulate these
CC   activities. The PAR4 protein is further useful for dissecting the
CC   effects of thrombin or other activating proteases in the clotting
CC   pathway from the effects of these proteases at the cellular level.
CC   Agonists are specifically useful in promoting the proliferation
CC   and/or differentiation of platelets, in mediating inflammatory events,
CC   responses to vascular injury, chemotaxis or mitogenesis, and in
CC   producing growth factors. Antagonists are useful as research reagents
CC   for characterising sites of ligand-receptor interaction.
XX
SQ   Sequence 4895 BP; 971 A; 1573 C; 1400 G; 951 T; 0 other;
```

```
Alignment Scores:
Pred. No.:              3.92e-38        Length:         4895
Score:                  523.50          Matches:        121
Percent Similarity:     53.57%          Conservative:   59
Best Local Similarity:  36.01%          Mismatches:     145
Query Match:            24.27%          Indels:         12
DB:                     20              Gaps:           5

US-09-208-629F-6 (1-408) x AAZ32747 (1-4895)

Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
              ||||||   |||      |||      |||   |||   |||||||||
Db        225 CTGGCGGCACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGA-GGTGGTGAT 283

Qy         68 GluGluPhePro------PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85
              :::         |||         |||      |||:::   |||
Db        284 GACAGCACGCCCTCAATCCTGCCTGCCCCCCGCGGCTACCCAGGCCAAGTC--------- 334
```

```
Qy          86 LysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105
Db         335 ---------TGTGCCAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGGCA 382

Qy         106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
Db         383 CTGCTTCTGGGCTGGGTGCCCACCAGGCTGGTGCCCGCCCTCTATGGGCTGGTCCTGGTG 442

Qy         126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db         443 GTGGGGCTGCCGGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACGCAGGCACCTCGGCTG 502

Qy         146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
Db         503 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCCTGCTGGCCCTGGCGCTG 562

Qy         166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
Db         563 CCCCCGCGGATCGCCTACCACCTGCGTGGCCAGCGCTGGCCCTTCGGGGAGGCCGCCTGC 622

Qy         186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205
Db         623 CGCCTGGCCACGGCCGCACTCTATGGTCACATGTATGGCTCAGTGCTGCTGCTGGCCGCC 682

Qy         206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
Db         683 GTCAGCCTGGATCGCTACCTGGCCCTGGTGCACCCGCTGCGGGCCCGCGCCCTGCGTGGC 742

Qy         226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
Db         743 CGGCGCCTGGCCCTTGGACTCTGCATGGCTGCTTGGCTCATGGCGGCCGCCCTGGCACTG 802

Qy         246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
Db         803 CCCCTGACACTGCAGCGGCAGACCTTCCGGCTGGCGCGCTCCGATCGCGTGCTCTGCCAT 862

Qy         266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
Db         863 GACGCGCTGCCCCTGGACGCACAGGCCTCCCACTGGCAA---CCGGCCTTCACCTGCCTG 919

Qy         286 AlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305
Db         920 GCGCTGTTGGGCTGTTTCCTGCCCCTGCTGGCCATGCTGCTGTGCTACGGGGCCACCCTG 979

Qy         306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325
Db         980 CACACGCTGGCGGCCAGCGGCCGGCGCTACGGCCACGCGCTGAGGCTGACCGCAGTGGTG 1039

Qy         326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345
Db        1040 CTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGCAACCTGCTGCTGCTGCTGCATTACTCG 1099

Qy         346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365
Db        1100 GACCCGAGCCCCAGCGCCTGGGGCAACCTCTATGGTGCCTACGTGCCCAGCCTGGCGCTG 1159

Qy         366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db        1160 AGCACCCTCAACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1207
```

RESULT 22
AAZ50775
ID AAZ50775 standard; DNA; 4895 BP.
XX
AC AAZ50775;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human protease activated receptor-4 DNA.
XX
KW Human; PAR-4; protease activated receptor;
KW antisense molecule; PAR antibody; cytostatic; therapeutic;
KW metastatic tumour cell; placental implantation; invasive cell; ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   176..1333
FT   /*tag= a
FT   /product= "PAR-4"
XX
PN WO200008150-A1.
XX
PD 17-FEB-2000.
XX
PF 05-FEB-1999; 99WO-IL00079.
XX
PR 07-AUG-1998; 98IL-0125698.
XX
PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
PI Bar-Shavit R;
XX
DR WPI; 2000-205706/18.
DR P-PSDB; AAY45036.
XX
PT Treating metastatic tumor cells useful for treating disorders involving
PT placenta implantation in a female comprises administration of an
PT antisense molecule complementary to an RNA sequence of a protease
PT activated receptor protein -
XX
PS Example 3; Fig 11a; 46pp; English.
XX
CC The patent discloses a method to treat metastatic tumour cells using
CC an antisense molecule comprising a polynucleotide complementary to an
CC RNA sequence of a protease activated receptor (PAR) protein, or an
CC antibody capable of binding to a PAR protein. The antisense molecules and
CC antibodies of PAR protein are also used to treat disorders associated
CC with implantation of placenta. The present sequence is a human PAR-4
CC DNA used for producing antisense molecules for treating invasive
CC cells.
XX
SQ Sequence 4895 BP; 971 A; 1573 C; 1400 G; 951 T; 0 other;

Alignment Scores:

Pred. No.:	3.92e-38	Length:	4895
Score:	523.50	Matches:	121
Percent Similarity:	53.57%	Conservative:	59
Best Local Similarity:	36.01%	Mismatches:	145
Query Match:	24.27%	Indels:	12
DB:	21	Gaps:	5

US-09-208-629F-6 (1-408) x AAZ50775 (1-4895)

```
Qy          49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
Db         225 CTGGCGGCACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGA-GGTGGTGAT 283

Qy          68 GluGluPhePro------PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85
Db         284 GACAGCACGCCCTCAATCCTGCCTGCCCCCCGCGGCTACCCAGGCCAAGTC--------- 334

Qy          86 LysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105
Db         335 ---------TGTGCCAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGGCA 382

Qy         106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
Db         383 CTGCTTCTGGGCTGGGTGCCCACCAGGCTGGTGCCCGCCCTCTATGGGCTGGTCCTGGTG 442

Qy         126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db         443 GTGGGGCTGCCGGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACGCAGGCACCTCGGCTG 502

Qy         146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
```

```
Db        503 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCCTGCTGGCCCTGGCGCTG 562

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185

Db        563 CCCCCGCGGATCGCCTACCACCTGCGTGGCCAGCGCTGGCCCTTCGGGGAGGCCGCCTGC 622

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205

Db        623 CGCCTGGCCACGGCCGCACTCTATGGTCACATGTATGGCTCAGTGCTGCTGCTGGCCGCC 682

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225

Db        683 GTCAGCCTGGATCGCTACCTGGCCCTGGTGCACCCGCTGCGGGCCCGCGCCCTGCGTGGC 742

Qy        226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245

Db        743 CGGCGCCTGGCCCTTGGACTCTGCATGGCTGCTTGGCTCATGGCGGCCGCCCTGGCACTG 802

Qy        246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265

Db        803 CCCCTGACACTGCAGCGGCAGACCTTCCGGCTGGCGCGCTCCGATCGCGTGCTCTGCCAT 862

Qy        266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285

Db        863 GACGCGCTGCCCCTGGACGCACAGGCCTCCCACTGGCAA---CCGGCCTTCACCTGCCTG 919

Qy        286 AlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305

Db        920 GCGCTGTTGGGCTGTTTCCTGCCCCTGCTGGCCATGCTGCTGTGCTACGGGGCCACCCTG 979

Qy        306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325

Db        980 CACACGCTGGCGGCCAGCGGCCGGCGCTACGGCCACGCGCTGAGGCTGACCGCAGTGGTG 1039

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345

Db       1040 CTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGCAACCTGCTGCTGCTGCTGCATTACTCG 1099

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365

Db       1100 GACCCGAGCCCCAGCGCCTGGGGCAACCTCTATGGTGCCTACGTGCCCAGCCTGGCGCTG 1159

Qy        366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381

Db       1160 AGCACCCTCAACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1207

RESULT 23
AAH43632
ID   AAH43632 standard; cDNA; 4925 BP.
XX
AC   AAH43632;
XX
DT   21-JAN-2002  (first entry)
XX
DE   Human PAR4 coding sequence.
XX
KW   Human; protease-activated receptor; PAR4; N-terminal; activated PAR4;
KW   platelet aggregation; inhibition; tumour cell; proliferation; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             191..1348
FT                   /*tag=  a
FT                   /product= "PAR4"
FT                   /note= "CDS derived from protein AAB47623"
XX
PN   WO200158930-A1.
XX
PD   16-AUG-2001.
XX
PF   06-FEB-2001; 2001WO-US03807.
XX
PR   09-FEB-2000; 2000US-0500646.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Baindur N,  West RR;
XX
DR   WPI; 2001-656678/75.
DR   P-PSDB; AAB47623.
XX
PT   Peptides comprising an amino acid sequence  are capable of stimulating
PT   protease-activated receptor 4 which are useful in diagnosis and therapy
PT   e.g. inhibiting tumor cell proliferation and stimulating platelet
PT   aggregation -
XX
PS   Disclosure; Page 51-58; 84pp; English.
XX
CC   This sequence encodes human protease-activated receptor, PAR4. Peptides
CC   derived from, or based on, the N-terminal of activated PAR4 (see
CC   AAB47624-77), are capable of stimulating PAR4.  These peptides may be
CC   used to activate PAR4 at lower concentrations than wild type PAR4.
CC   These peptides may be used to stimulate platelet aggregation, and for
CC   inhibiting tumor cell proliferation.
XX
SQ   Sequence 4925 BP; 977 A; 1582 C; 1411 G; 955 T; 0 other;

Alignment Scores:
Pred. No.:                3.95e-38      Length:          4925
Score:                    523.50        Matches:         121
Percent Similarity:       53.57%        Conservative:    59
Best Local Similarity:    36.01%        Mismatches:      145
Query Match:              24.27%        Indels:          12
DB:                       22            Gaps:            5

US-09-208-629F-6 (1-408) x AAH43632 (1-4925)

Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67

Db        240 CTGGCGGCACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGA-GGTGGTGAT 298

Qy         68 GluGluPhePro------PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85

Db        299 GACAGCACGCCCTCAATCCTGCCTGCCCCCCGCGGCTACCCAGGCCAAGTC--------- 349

Qy         86 LysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105

Db        350 ---------TGTGCCAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGGCA 397

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125

Db        398 CTGCTTCTGGGCTGGGTGCCCACCAGGCTGGTGCCCGCCCTCTATGGGCTGGTCCTGGTG 457

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145

Db        458 GTGGGGCTGCCGGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACGCAGGCACCTCGGCTG 517

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165

Db        518 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCCTGCTGGCCCTGGCGCTG 577

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185

Db        578 CCCCCGCGGATCGCCTACCACCTGCGTGGCCAGCGCTGGCCCTTCGGGGAGGCCGCCTGC 637

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205

Db        638 CGCCTGGCCACGGCCGCACTCTATGGTCACATGTATGGCTCAGTGCTGCTGCTGGCCGCC 697

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225

Db        698 GTCAGCCTGGATCGCTACCTGGCCCTGGTGCACCCGCTGCGGGCCCGCGCCCTGCGTGGC 757

Qy        226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
```

```
Db        758 CGGCGCCTGGCCCTTGGACTCTGCATGGCTGCTTGGCTCATGGCGGCCGCCCTGGCACTG 817

Qy        246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265

Db        818 CCCCTGACACTGCAGCGGCAGACCTTCCGGCTGGCGCGCTCCGATCGCGTGCTCTGCCAT 877

Qy        266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285

Db        878 GACGCGCTGCCCCTGGACGCACAGGCCTCCCACTGGCAA---CCGGCCTTCACCTGCCTG 934

Qy        286 AlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305

Db        935 GCGCTGTTGGGCTGTTTCCTGCCCCTGCTGGCCATGCTGCTGTGCTACGGGGCCACCCTG 994

Qy        306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325

Db        995 CACACGCTGGCGGCCAGCGGCCGGCGCTACGGCCACGCGCTGAGGCTGACCGCAGTGGTG 1054

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345

Db       1055 CTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGCAACCTGCTGCTGCTGCTGCATTACTCG 1114

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365

Db       1115 GACCCGAGCCCCAGCGCCTGGGGCAACCTCTATGGTGCCTACGTGCCCAGCCTGGCGCTG 1174

Qy        366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381

Db       1175 AGCACCCTCAACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1222
```

RESULT 24
AAX90981
ID AAX90981 standard; cDNA; 1534 BP.
XX
AC AAX90981;
XX
DT 25-JAN-2000 (first entry)
XX
DE Human protease-activated receptor 4 (PAR4) cDNA.
XX
KW Human protease-activated receptor 4 cDNA; PAR4; thrombin receptor;
KW phosphoinositide hydrolysis; calcium efflux; platelet aggregation;
KW platelet activation; thrombin agonist; thrombin antagonist; therapeutic;
KW wound; blood coagulation; heart attack; stroke; inflammatory response;
KW proliferative response; atherosclerosis; restenosis; glomerulosclerosis;
KW pulmonary inflammation; ARDS; adult respiratory distress syndrome; ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   3..1160
FT   /*tag= a
FT   /product= "PAR4"
XX
PN WO9943809-A2.
XX
PD 02-SEP-1999.
XX
PF 11-FEB-1999; 99WO-US02983.
XX
PR 27-FEB-1998; 98US-0032397.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Coughlin SR, Kahn M;
XX
DR WPI; 1999-619953/53.
DR P-PSDB; AAY15082.
XX
PT DNA molecules encoding protease-activated receptor 4, useful in
PT compound assays for thrombin agonist and antagonist activity -
XX
PS Claim 3; Fig 4; 69pp; English.
XX
CC The present sequence is a cDNA encoding human protease-activated
CC receptor 4 (PAR4) which is a thrombin receptor expressed on cell surface.
CC PAR4 is activated by thrombin and mediates signalling events e.g.
CC phosphoinositide hydrolysis, calcium efflux and platelet aggregation.
CC The receptor is highly expressed in spleen cells and likely to be
CC involved in thrombin-mediated activation of platelets and other
CC haematopoietic cells. It is used for screening novel thrombin agonists
CC or antagonists. The agonists are used as therapeutics to treat wounds,
CC promote clotting and as reagents to activate platelets in diagnostic
CC tests. Antagonists are used to control blood coagulation, treat heart
CC attacks and strokes, and block inflammatory and proliferative responses
CC that occur in normal wound healing and variety of diseases including
CC atherosclerosis, restenosis, pulmonary inflammation (ARDS) and
CC glomerulosclerosis.
XX
SQ Sequence 1534 BP; 235 A; 526 C; 479 G; 293 T; 1 other;

Alignment Scores:

Pred. No.:	9.75e-39	Length:	1534
Score:	523.00	Matches:	121
Percent Similarity:	53.57%	Conservative:	59
Best Local Similarity:	36.01%	Mismatches:	145
Query Match:	24.25%	Indels:	12
DB:	20	Gaps:	5

US-09-208-629F-6 (1-408) x AAX90981 (1-1534)

```
Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67

Db         52 CTGGCGGCACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGA-GGTGGTGAT 110

Qy         68 GluGluPhePro------PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85

Db        111 GACAGCACGCCCTCAATCCTGCCTGCCCCCCGCGGCTACCCAGGCCAAGTC--------- 161

Qy         86 LysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105

Db        162 ---------TGTGCCAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGGCA 209

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125

Db        210 CTGCTTCTGGGCTGGGTGCCCACCAGGCTGGTGCCCGCCCTCTATGGGCTGGTCCTGGTG 269

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145

Db        270 GTGGGGCTGCCGGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACGCAGGCACCTCGGCTG 329

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165

Db        330 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCCTGCTGGCCCTGGCGCTG 389

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185

Db        390 CCCCCGCGGATCGCCTACCACCTGCGTGGCCAGCGCTGGCCCTTCGGGGAGGCCGCCTGC 449

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205

Db        450 CGCCTGGCCACGGCCGCACTCTATGGTCACATGTATGGCTCAGTGCTGCTGCTGGCCGCC 509

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225

Db        510 GTCAGCCTGGATCGCTACCTGGCCCTGGTGCACCCGCTGCGGGCCCGCGCCCTGCGTGGC 569

Qy        226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245

Db        570 CGGCGCCTGGCCCTTGGACTCTGCATGGCTGCTTGGCTCATGGCGGCCGCCCTGGCACTG 629

Qy        246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265

Db        630 CCCCTGACACTGCAGCGGCAGACCTTCCGGCTGGCGCGCTCCGATCGCGTGCTCTGCCAT 689
```

```
Qy        266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
              |||                        :::|||   :::|||         |||      |||
Db        690 GACGCGCTGCCCCTGGACGCACAGGCCTCCCACTGGCAA---CCGGCCTTCACCTGCCTG 746

Qy        286 AlaPheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305
              |||      |||      :::|||   :::   ::::::   ||||||   |||   :::
Db        747 GCGCTGTTGGGCTGTTTCCTGCCCCTGCTGGCCATGCTGCTGTGCTACGGGGCCACCCTG 806

Qy        306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325
              ||||||   |||        |||:::        ::::::   :::   ::::::
Db        807 CACACGCTGGCGGCCAGCGGCCGGCGCTACGGCCACGCGCTGAGGCTGACCGCAGTGGTG 866

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345
              |||                    |||   |||||||||:::::::|||:::::::|||::::::
Db        867 CTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGCAACCTGCTGCTGCTGCTGCATTACTCG 926

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365
              :::      :::            ||||||      |||:::   :::|||   |||
Db        927 GACCCGAGCCCCAGCGCCTGGGGCAACCTCTATGGTGCCTACGTGCCCAGCCTGGCGCTG 986

Qy        366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
              :::||||||||||||:::|||||||||:::|||:::   :::|||
Db        987 AGCACCCTCAACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1034

RESULT 25
AAF21431
ID   AAF21431 standard; DNA; 1534 BP.
XX
AC   AAF21431;
XX
DT   14-MAR-2001  (first entry)
XX
DE   Human low adenosine antisense oligonucleotide related sequence #2998.
XX
KW   Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW   human; airway disorder; bronchoconstriction; lung inflammation;
KW   surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW   immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW   respiratory obstruction; pulmonary obstruction; impeded respiration;
KW   surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW   respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW   pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW   chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200062736-A2.
XX
PD   26-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US08020.
XX
PR   06-APR-1999;   99US-0127958.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
PA   (NYCE/) NYCE J W.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-679539/66.
XX
PT   Low adenosine (A) content antisense oligonucleotides which do not
PT   trigger adenosine receptors during metabolism, useful e.g. for treating
PT   cancers and respiratory obstructions -
XX
PS   Disclosure; Page 1415; 1592pp; English.
XX
CC   The present invention describes low adenosine (A) content antisense
CC   oligonucleotides and compositions (I) comprising them. In the antisense
CC   oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC   (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC   immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC   The antisense oligonucleotides and (I) can be used to down-regulate the
CC   expression and or activity of target polypeptides associated with
CC   lung/respiratory disorders and malignancies, such as stimulating and
CC   activating peptide factors and transmitters, transcription factors,
CC   immunoglobulins and antibodies, antibody receptors, cytokines and
CC   chemokines, endogenously produced specific and non-specific enzymes,
CC   binding proteins, adhesion molecules and their receptors, cytokine and
CC   chemokine receptors, adenosine receptors, bradykinin receptors, central
CC   nervous system (CNS) and peripheral nervous and non-nervous system
CC   receptors, CNS and peripheral nervous and non-nervous system peptide
CC   transmitters, defensins, growth factors, vasoactive peptides and
CC   receptors, binding proteins and malignancy associated proteins. The
CC   antisense oligonucleotides may be used in this way to treat disorders
CC   including respiratory obstruction (especially pulmonary obstruction
CC   and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC   and/or surfactant hypoproduction which are associated with a disease or
CC   condition selected from pulmonary vasoconstriction, inflammation,
CC   allergies, asthma, impeded respiration, respiratory distress syndrome
CC   (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC   hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC   pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC   and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC   fragments and antisense oligonucleotides used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 1534 BP; 235 A; 526 C; 479 G; 293 T; 1 other;

Alignment Scores:
Pred. No.:              9.75e-39        Length:         1534
Score:                  523.00          Matches:        121
Percent Similarity:     53.57%          Conservative:   59
Best Local Similarity:  36.01%          Mismatches:     145
Query Match:            24.25%          Indels:         12
DB:                     21              Gaps:           5

US-09-208-629F-6 (1-408) x AAF21431 (1-1534)

Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
              ||||||   |||      |||      |||   |||   ||||||||
Db         52 CTGGCGGCACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGA-GGTGGTGAT 110

Qy         68 GluGluPhePro------PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85
              :::      |||           |||      |||:::   |||
Db        111 GACAGCACGCCCTCAATCCTGCCTGCCCCCCGCGGCTACCCAGGCCAAGTC--------- 161

Qy         86 LysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105
                        |||      :::|||   :::   |||   :::   :::::::::
Db        162 ---------TGTGCCAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGGCA 209

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
              |||          :::   |||:::|||:::||||||:::|||   ||||||   |||
Db        210 CTGCTTCTGGGCTGGGTGCCCACCAGGCTGGTGCCCGCCCTCTATGGGCTGGTCCTGGTG 269

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
              ||||||:::|||||||||   :::   ||||||:::|||       :::          :::
Db        270 GTGGGGCTGCCGGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACGCAGGCACCTCGGCTG 329

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
                 :::|||:::          |||||||||   ||||||   |||      :::   |||
Db        330 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCCTGCTGGCCCTGGCGCTG 389

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
              |||   :::|||||||||||||||   |||      |||   |||||||||      |||
Db        390 CCCCCGCGGATCGCCTACCACCTGCGTGGCCAGCGCTGGCCCTTCGGGGAGGCCGCCTGC 449

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205
              |||      |||         ||||||:::||||||   |||:::||||||||||||
Db        450 CGCCTGGCCACGGCCGCACTCTATGGTCACATGTATGGCTCAGTGCTGCTGCTGGCCGCC 509

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
              :::|||::::::||||||||||||:::|||||||||          |||   |||
Db        510 GTCAGCCTGGATCGCTACCTGGCCCTGGTGCACCCGCTGCGGGCCCGCGCCCTGCGTGGC 569
```

```
Qy        226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
Db        570 CGGCGCCTGGCCCTTGGACTCTGCATGGCTGCTTGGCTCATGGCGGCCGCCCTGGCACTG 629

Qy        246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
Db        630 CCCCTGACACTGCAGCGGCAGACCTTCCGGCTGGCGCGCTCCGATCGCGTGCTCTGCCAT 689

Qy        266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
Db        690 GACGCGCTGCCCCTGGACGCACAGGCCTCCCACTGGCAA---CCGGCCTTCACCTGCCTG 746

Qy        286 AlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305
Db        747 GCGCTGTTGGGCTGTTTCCTGCCCCTGCTGGCCATGCTGCTGTGCTACGGGGCCACCCTG 806

Qy        306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325
Db        807 CACACGCTGGCGGCCAGCGGCCGGCGCTACGGCCACGCGCTGAGGCTGACCGCAGTGGTG 866

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345
Db        867 CTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGCAACCTGCTGCTGCTGCTGCATTACTCG 926

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365
Db        927 GACCCGAGCCCCAGCGCCTGGGGCAACCTCTATGGTGCCTACGTGCCCAGCCTGGCGCTG 986

Qy        366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db        987 AGCACCCTCAACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1034
```

RESULT 26
AAA35309
ID AAA35309 standard; DNA; 1534 BP.
XX
AC AAA35309;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:183.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -
XX
PS Disclosure; Page 1330; 1343pp; English.

XX
CC The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.
XX
SQ Sequence 1534 BP; 235 A; 526 C; 479 G; 293 T; 1 other;

Alignment Scores:

Pred. No.:	9.75e-39	Length:	1534
Score:	523.00	Matches:	121
Percent Similarity:	53.57%	Conservative:	59
Best Local Similarity:	36.01%	Mismatches:	145
Query Match:	24.25%	Indels:	12
DB:	21	Gaps:	5

US-09-208-629F-6 (1-408) x AAA35309 (1-1534)

```
Qy         49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
Db         52 CTGGCGGCACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGA-GGTGGTGAT 110

Qy         68 GluGluPhePro------PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85
Db        111 GACAGCACGCCCTCAATCCTGCCTGCCCCCCGCGGCTACCCAGGCCAAGTC--------- 161

Qy         86 LysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105
Db        162 ---------TGTGCCAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGGCA 209

Qy        106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
Db        210 CTGCTTCTGGGCTGGGTGCCCACCAGGCTGGTGCCCGCCCTCTATGGGCTGGTCCTGGTG 269

Qy        126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db        270 GTGGGGCTGCCGGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACGCAGGCACCTCGGCTG 329

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
Db        330 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCCTGCTGGCCCTGGCGCTG 389

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
Db        390 CCCCCGCGGATCGCCTACCACCTGCGTGGCCAGCGCTGGCCCTTCGGGGAGGCCGCCTGC 449

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205
Db        450 CGCCTGGCCACGGCCGCACTCTATGGTCACATGTATGGCTCAGTGCTGCTGCTGGCCGCC 509

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
```

```
Db        510 GTCAGCCTGGATCGCTACCTGGCCCTGGTGCACCCGCTGCGGGCCCGCGCCCTGCGTGGC 569

Qy        226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245

Db        570 CGGCGCCTGGCCCTTGGACTCTGCATGGCTGCTTGGCTCATGGCGGCCGCCCTGGCACTG 629

Qy        246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265

Db        630 CCCCTGACACTGCAGCGGCAGACCTTCCGGCTGGCGCGCTCCGATCGCGTGCTCTGCCAT 689

Qy        266 AspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285

Db        690 GACGCGCTGCCCCTGGACGCACAGGCCTCCCACTGGCAA---CCGGCCTTCACCTGCCTG 746

Qy        286 AlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIle 305

Db        747 GCGCTGTTGGGCTGTTTCCTGCCCCTGCTGGCCATGCTGCTGTGCTACGGGGCCACCCTG 806

Qy        306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325

Db        807 CACACGCTGGCGGCCAGCGGCCGGCGCTACGGCCACGCGCTGAGGCTGACCGCAGTGGTG 866

Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345

Db        867 CTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGCAACCTGCTGCTGCTGCTGCATTACTCG 926

Qy        346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365

Db        927 GACCCGAGCCCCAGCGCCTGGGGCAACCTCTATGGTGCCTACGTGCCCAGCCTGGCGCTG 986

Qy        366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381

Db        987 AGCACCCTCAACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1034

RESULT 27
AAZ32191
ID   AAZ32191 standard; cDNA; 3472 BP.
XX
AC   AAZ32191;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Human thrombin receptor nucleotide sequence.
XX
KW   Human; coding sequence polymorphism; vascular pathology gene;
KW   polymorphic site; phenotype correlation; forensic; paternity testing;
KW   medicine; genetic analysis; vascular disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO9950454-A2.
XX
PD   07-OCT-1999.
XX
PF   26-MAR-1999;   99WO-US06473.
XX
PR   01-APR-1998;   98US-0054272.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Lander ES,  Daley GQ,  Cargill M,  Ireland JS,  Rozen SG;
XX
DR   WPI; 1999-620066/53.
DR   P-PSDB; AAY49570.
XX
PT   Determination of polymorphisms in genes, especially those identifying
PT   predisposition to vascular disease  -
XX
PS   Claim 1; Fig 35; 134pp; English.
XX
CC   AAZ32159 to AAZ32194 represent reference alleles for specifically
CC   claimed nucleic acid sequences from the present invention which comprise
CC   polymorphic sites as given in a table in the specification, selected
CC   from 92 single nucleotide polymorphisms in which the nucleotide at the
CC   polymorphic site is different from a nucleotide at the same site in a
CC   reference allele. The nucleic acids, and primers and probes, are used to
CC   identify polymorphisms, which may predispose an individual to disease,
CC   especially a vascular disease. They can also be used in phenotype
CC   correlations, forensics, paternity testing, medicine or genetic
CC   analysis. AAY49550 to AAY49573 represent the proteins which correspond
CC   to some of the reference alleles.
XX
SQ   Sequence 3472 BP; 933 A; 817 C; 785 G; 937 T; 0 other;

Alignment Scores:
Pred. No.:                2.8e-38       Length:          3472
Score:                    523.00        Matches:         127
Percent Similarity:       52.59%        Conservative:    76
Best Local Similarity:    32.90%        Mismatches:      153
Query Match:              24.25%        Indels:          30
DB:                       20            Gaps:            11

US-09-208-629F-6 (1-408) x AAZ32191 (1-3472)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33

Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53

Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72

Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----------IleThrValLysIleLys 88

Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108

Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128

Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146

Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166

Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186

Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206

Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223

Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243

Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263

Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983
```

```
Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||              :::      |||||||||
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::          |||   :::|||   ::::::         ||||||   :::
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              |||||||||   |||:::                  :::          |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::           |||      :::      ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||::::::||||||   |||::::::         :::   :::   |||:::
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::||||||
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1332 ATTTACTATTACGCTTCC 1349

RESULT 28
AAZ50771
ID   AAZ50771 standard; DNA; 3480 BP.
XX
AC   AAZ50771;
XX
DT   31-MAY-2000  (first entry)
XX
DE   Human thrombin receptor DNA.
XX
KW   Human; thrombin receptor; ThR; PAR-1; protease activated receptor;
KW   antisense molecule; PAR antibody; cytostatic; therapeutic;
KW   metastatic tumour cell; placental implantation; invasive cell; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1502
FT                   /*tag=  a
FT                   /product= "Thrombin receptor"
FT                   /transl_except= (pos:1500..1502, aa:Glx)
FT                   /note= "no stop codon given"
FT   misc_binding    320..570
FT                   /*tag=  b
FT                   /bound_moiety= "ThR RNA probe"
XX
PN   WO200008150-A1.
XX
PD   17-FEB-2000.
XX
PF   05-FEB-1999;   99WO-IL00079.
XX
PR   07-AUG-1998;   98IL-0125698.
XX
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
PI   Bar-Shavit R;
XX
DR   WPI; 2000-205706/18.
DR   P-PSDB; AAY45035.
XX
PT   Treating metastatic tumor cells useful for treating disorders involving
PT   placenta implantation in a female comprises administration of an
PT   antisense molecule complementary to an RNA sequence of a protease
PT   activated receptor protein -
XX
PS   Example 2; Fig 1a; 46pp; English.
XX
CC   The patent discloses a method to treat metastatic tumour cells using
CC   an antisense molecule comprising a polynucleotide complementary to an
CC   RNA sequence of a protease activated receptor (PAR) protein, or an
CC   antibody capable of binding to a PAR protein. The antisense molecules and
CC   antibodies of PAR protein are also used to treat disorders associated
CC   with implantation of placenta. The present sequence is a DNA encoding
CC   human thrombin receptor (ThR) (also known as PAR-1), which is a
CC   G-coupled protein belonging to the PAR family. This sequence is used to
CC   produce antisense molecules or antibodies against PAR protein for
CC   therapeutic use.
XX
SQ   Sequence 3480 BP; 941 A; 817 C; 785 G; 937 T; 0 other;

Alignment Scores:
Pred. No.:                2.81e-38      Length:          3480
Score:                    523.00        Matches:         127
Percent Similarity:       52.59%        Conservative:    76
Best Local Similarity:    32.90%        Mismatches:      153
Query Match:              24.25%        Indels:          30
DB:                       21            Gaps:            11

US-09-208-629F-6 (1-408) x AAZ50771 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||      .|||:::      |||   ||||||
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                  |||      :::                  ||||||                |||
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||      ::::::|||             ||||||      :::|||   |||
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----------IleThrValLysIleLys 88
                    :::      :::|||   |||           :::::::::
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
              |||   :::                 :::                ||||||||||
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||      :::|||:::::|||      ||||||||||||   :::
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||      |||:::   :::::::::|||      :::   :::
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::      :::||||||   ||||||   ||||||          ||||||
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||::::::|||   ||||||      |||||||||
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                 |||      ||||||   |||||||||   ||||||||||||:::      |||
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||           ::::::|||   |||
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
              :::   ::::::            ||||||      :::||||||
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923
```

```
Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ::::::|||      :::   :::|||         :::         :::|||||||||
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                 :::     ||||||||
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              :::   :::         |||   :::|||   ::::::         ||||||   :::
Db       1041 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
              ||||||||   |||:::                     :::         |||   |||
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              :::         |||      :::      ||||||   |||||||||   |||:::
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
              |||::::::||||||   |||::::::         :::   :::   |||:::
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::|||||
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1332 ATTTACTATTACGCTTCC 1349
```

RESULT 29
ABK70889
ID ABK70889 standard; cDNA; 1278 BP.
XX
AC ABK70889;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cDNA encoding PAR1 type thrombin receptor.
XX
KW Human; ss; gene; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.
XX
OS Homo sapiens.
XX
PN JP2002010784-A.
XX
PD 15-JAN-2002.
XX
PF 29-JUN-2000; 2000JP-0196514.
XX
PR 29-JUN-2000; 2000JP-0196514.
XX
PA (TEIJ) TEIJIN LTD.
XX
DR WPI; 2002-321520/36.
DR P-PSDB; ABG35300.
XX
PT An inhibitor of cell growth mediated by thrombin used to treat
PT inflammatory and cell proliferative diseases -
XX
PS Claim 9; Page 26-27; 44pp; Japanese.
XX
CC The invention relates to a polypeptide or a compound which can inhibit
CC cell growth caused by thrombin. The polypeptide/compound combines to a
CC specific region of the structure of PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a cDNA encoding a PAR1 type thrombin receptor (or a
CC modified version).
XX
SQ Sequence 1278 BP; 276 A; 353 C; 284 G; 365 T; 0 other;

Alignment Scores:

Pred. No.:	1.8e-38	Length:	1278
Score:	519.00	Matches:	126
Percent Similarity:	52.33%	Conservative:	76
Best Local Similarity:	32.64%	Mismatches:	154
Query Match:	24.06%	Indels:	30
DB:	24	Gaps:	11

US-09-208-629F-6 (1-408) x ABK70889 (1-1278)

```
Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||      |||:::      |||   ||||||
Db          4 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 63

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||      :::                  ||||||                |||
Db         64 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT------------GCCACC 111

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||      ::::::|||             ||||||      :::|||   |||
Db        112 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 171

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr------------IleThrValLysIleLys 88
                    :::      :::|||   |||            :::::::::
Db        172 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 231

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
                 |||   :::                 :::                ||||||||||||
Db        232 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 291

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||       :::|||::::::|||      ||||||||||||   :::
Db        292 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 351

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||       |||:::   :::::::::|||       :::   :::
Db        352 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 408

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::      :::||||||   ||||||   ||||||            ||||||
Db        409 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 468

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||:::::|||   ||||||      |||||||||
Db        469 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 528

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                 |||      ||||||   |||||||||   |||||||||||:::      |||
Db        529 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 588

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||           ::::::|||   |||
Db        589 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 648

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
                 :::   ::::::         ||||||      :::|||||
Db        649 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 699

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ::::::|||      :::   :::|||         :::         :::|||||||||
Db        700 GTAGTGCCTCTGCTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 759

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   |||                 :::     ||||||||
```

```
Db        760 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 816

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303

Db        817 GCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 876

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316

Db        877 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 936

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336

Db        937 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA 987

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355

Db        988 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1047

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375

Db       1048 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTGCTGCATCGACCCCCTA 1107

Qy        376 LeuTyrPheLeuMetSer 381

Db       1108 ATTTACTATTACGCTTCC 1125

RESULT 30
AAQ73590
ID   AAQ73590 standard; DNA; 1764 BP.
XX
AC   AAQ73590;
XX
DT   25-JUN-1995  (first entry)
XX
DE   Fragment of the human thrombin receptor gene.
XX
KW   TR; expression; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             376..1653
FT                   /*tag=  a
XX
PN   WO9421789-A.
XX
PD   29-SEP-1994.
XX
PF   28-FEB-1994;   94WO-US02388.
XX
PR   25-MAR-1993;   93US-0038662.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Betlach MC,  Turner GJ;
XX
DR   WPI; 1994-317010/39.
DR   P-PSDB; AAR60698.
XX
PT   Expression of heterologous proteins in halo-bacteria - using
PT   regulatory and stop sequences from halo-bacteria, pref. the
PT   bacterio-rhodopsin gene.
XX
PS   Disclosure; Fig 14; 118pp; English.
XX
CC   The sequence is that of the human thrombin receptor gene fragment.
CC   The gene is used to exemplify a new expression vector for producing
CC   heterologous polypeptides in a halobacterial host.
CC   See also AAQ73586-92.
XX
SQ   Sequence 1764 BP; 382 A; 492 C; 421 G; 469 T; 0 other;

Alignment Scores:

Pred. No.:              4.65e-38        Length:         1764
Score:                  516.50          Matches:        123
Percent Similarity:     53.05%          Conservative:   77
Best Local Similarity:  32.63%          Mismatches:     144
Query Match:            23.95%          Indels:         33
DB:                     15              Gaps:           11

US-09-208-629F-6 (1-408) x AAQ73590 (1-1764)

Qy         30 LeuLeuLeuLeuLeuProThrPheCysGln--------------------SerGlyMet 42

Db        376 ATGTTGGAGTTATTGCCAACAGCAGTGGAGGGGGTATCGCAGGCCCAGATCCAGGCGCTG 435

Qy         43 GluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAla 62

Db        436 GACTACAAGGACGATGATGACGTCGACGCCACCTTAGATCCCCGGTCATTTCTTCTCAGG 495

Qy         63 ProProAsn---SerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81

Db        496 AACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAGAAAAATGAAAGTGGGTTA 555

Qy         82 Thr------------IleThrValLysIleLysCysProGluGluSerAlaSerHisLeu 97

Db        556 ACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTTCAAAAACAACTTCCTGCA 615

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117

Db        616 TTCATCTCAGAAGATGCCTCCGGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCA 675

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsn------AlaValThrLeu 135

Db        676 TCTGTGTACACCGGAGTGTTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTG 735

Qy        136 TrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAla 155

Db        736 TTCATCCTGAAAATGAAGGTCAAGAAG---CCGGCGGTGGTGTACATGCTGCACCTGGCC 792

Qy        156 IleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGly 175

Db        793 ACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCCTTTAAGATCAGCTATTACTTTTCCGGC 852

Qy        176 AsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsn 195

Db        853 AGTGATTGGCAGTTTGGGTCTGAATTGTGTCGCTTCGTCACTGCAGCATTTTACTGTAAC 912

Qy        196 MetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleVal 215

Db        913 ATGTACGCCTCTATCTTGCTCATGACAGTCATAAGCATTGACCGGTTTCTGGCTGTGGTG 972

Qy        216 HisPro---------PheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThr 232

Db        973 TATCCCATGCAGTCCCTCTCCTGGCGTACTCTGGGAAGGGCTTCCTTC---------ACT 1023

Qy        233 CysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGln 252

Db       1024 TGTCTGGCCATCTGGGCTTTGGCCATCGCAGGGGTAGTGCCTCTCGTCCTCAAGGAGCAA 1083

Qy        253 GluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGlu 272

Db       1084 ACCATCCAGGTGCCCGGGCTCAACATCACTACCTGTCATGATGTGCTCAATGAAACCCTG 1143

Qy        273 SerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292

Db       1144 CTCGAAGGC---TACTATGCCTACTACTTCTCAGCCTTCTCTGCTGTCTTCTTTTTTGTG 1200

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsn--------- 309

Db       1201 CCGCTGATCATTTCCACGGTCTGTTATGTGTCTATCATTCGATGTCTTAGCTCTTCCGCA 1260

Qy        310 -----------AlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325

Db       1261 GTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTGTTC---------CTGTCAGCTGCTGTT 1311
```

```
Qy        326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAla 345
Db       1312 TTCTGCATCTTCATCATTTGCTTCGGACCCACAAACGTCCTCCTGATTGCGCATTACTCA 1371

Qy        346 AsnTyr---TyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCys 364
Db       1372 TTCCTTTCTCACACTTCCACCACAGAGGCTGCCTACTTTGCCTACCTCCTCTGTGTCTGT 1431

Qy        365 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db       1432 GTCAGCAGCATAAGCTCGTGCATCGACCCCCTAATTTACTATTACGCTTCC 1482
```

RESULT 31
AAT62461
ID AAT62461 standard; DNA; 2910 BP.
XX
AC AAT62461;
XX
DT 16-AUG-1997 (first entry)
XX
DE ThR-GPA1 fusion gene.
XX
KW G-protein coupled receptor; agonist; antagonist; assay;
KW G-alpha protein; Gpa1p; GPA1 gene; thrombin receptor; ds.
XX
OS Chimaeric Homo sapiens;
OS Chimaeric Saccharomyces cerevisiae.
XX

FH	Key	Location/Qualifiers
FT	CDS	1..2679
FT		/*tag= a
FT		/product= ThR-GPA1 fusion
FT	mRNA	1..39
FT		/*tag= b
FT		/note= "STET nucleotides 535-573"
FT	mRNA	40..1251
FT		/*tag= c
FT		/product= thrombin receptor
FT		/note= "corresponds to nucleotides 288-1499 of the thrombin receptor gene"
FT	mRNA	1252..1260
FT		/*tag= d
FT		/note= "STET nucleotides 1827..1835"
FT	mRNA	1261..2910
FT		/*tag= e
FT		/product= Gpa1p
FT		/note= "corresponds to nucleotides 202-1851 of GPA1"
FT	primer_bind	complement (42..62)
FT		/*tag= f
FT		/note= "oligonucleotide g for thrombin receptor gene amplification"
FT	primer_bind	1229..1251
FT		/*tag= g
FT		/note= "oligonucleotide h for thrombin receptor gene amplification"

XX
PN WO9711159-A1.
XX
PD 27-MAR-1997.
XX
PF 20-SEP-1996; 96WO-US15203.
XX
PR 20-SEP-1995; 95US-0004023.
XX
PA (HEAR-) HEARTLAND BIOTECHNOLOGIES LLC.
XX
PI Boulton TG, Das PO, Mandell RB, McMullen TW;
XX
DR WPI; 1997-202868/18.
DR P-PSDB; AAW16314.
XX
PT New yeast constructs for use in screening assays - which express a
PT fusion comprising a mammalian or fungal transmembrane receptor and
PT a yeast G-alpha protein
XX
PS Example 2; Fig 12A-G; 109pp; English.
XX
CC A gene fusion (AAT62461) codes for a polypeptide (AAW16314) comprising
CC the human thrombin receptor (ThR) covalently linked to yeast G-alpha
CC protein Gpa1p (encoded by the GPA1 gene). It was obtd. by PCR
CC amplification (see also AAT62466-67) of the ThR gene and insertion
CC of the amplified sequence into pRMHBT3, which contains the GPA1
CC gene. Thrombin-dependent activation of the yeast mating pathway
CC by the ThR-GPA1 fusion protein was observed in ste2 gpa1- yeast
CC cells. The polypeptide is an example of novel fusion proteins
CC between eukaryotic heterotrimeric G-protein coupled receptors and
CC yeast G-alpha proteins that can be coupled to the pheromone-
CC induced signal transduction pathway of yeast and used in screening
CC assays to identify agonists or antagonists of the receptor.
XX
SQ Sequence 2910 BP; 868 A; 606 C; 628 G; 808 T; 0 other;

Alignment Scores:

Pred. No.:	2.31e-37	Length:	2910
Score:	512.00	Matches:	118
Percent Similarity:	54.47%	Conservative:	71
Best Local Similarity:	34.01%	Mismatches:	133
Query Match:	23.74%	Indels:	26
DB:	18	Gaps:	9

US-09-208-629F-6 (1-408) x AAT62461 (1-2910)

```
Qy         52 ProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPhePro 71
Db         86 CCTTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAAT-GATAAATATGAACCATTTTGG 144

Qy         72 PheSerAlaLeuGluGlyTrpThrGlyAlaThr------------IleThrValLysIle 87
Db        145 GAGGATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAA 204

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db        205 AGCAGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTG 264

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
Db        265 ACCAGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGC 324

Qy        128 ValProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db        325 CTCCCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG--- 381

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
Db        382 CCGGCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTC 441

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
Db        442 CCCTTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGT 501

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205
Db        502 CGCTTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTC 561

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGly 222
Db        562 ATAAGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACT 621

Qy        223 LeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeu 242
Db        622 CTGGGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCA 672

Qy        243 TyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThr 262
Db        673 GGGGTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACT 732
```

```
Qy        263 ThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPhe 282
Db        733 ACCTGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTC 789

Qy        283 IleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAla 302
Db        790 TCAGCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTG 849

Qy        303 AlaIleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrp 315
Db        850 TCTATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCT 909

Qy        316 LeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 335
Db        910 TTGTTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCC 960

Qy        336 SerAsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGly 354
Db        961 ACAAACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCT 1020

Qy        355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
Db       1021 GCCTACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCC 1080

Qy        375 PheLeuTyrPheLeuMetSer 381
Db       1081 CTAATTTACTATTACGCTTCC 1101
```

RESULT 32
AAF21432
ID AAF21432 standard; DNA; 3182 BP.
XX
AC AAF21432;
XX
DT 14-MAR-2001 (first entry)
XX
DE Human low adenosine antisense oligonucleotide related sequence #2999.
XX
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200062736-A2.
XX
PD 26-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US08020.
XX
PR 06-APR-1999; 99US-0127958.
XX
PA (UYEC-) UNIV EAST CAROLINA.
PA (NYCE/) NYCE J W.
XX
PI Nyce JW;
XX
DR WPI; 2000-679539/66.
XX
PT Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
PS Disclosure; Page 1415-1416; 1592pp; English.
XX
CC The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.
XX
SQ Sequence 3182 BP; 895 A; 701 C; 684 G; 902 T; 0 other;

Alignment Scores:

Pred. No.:	2.6e-37	Length:	3182
Score:	512.00	Matches:	118
Percent Similarity:	54.47%	Conservative:	71
Best Local Similarity:	34.01%	Mismatches:	133
Query Match:	23.74%	Indels:	26
DB:	21	Gaps:	9

US-09-208-629F-6 (1-408) x AAF21432 (1-3182)

```
Qy         52 ProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPhePro 71
Db         44 CCTTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAAT-GATAAATATGAACCATTTTGG 102

Qy         72 PheSerAlaLeuGluGlyTrpThrGlyAlaThr------------IleThrValLysIle 87
Db        103 GAGGATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAA 162

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db        163 AGCAGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTG 222

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
Db        223 ACCAGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGC 282

Qy        128 ValProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db        283 CTCCCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG--- 339

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
Db        340 CCGGCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTC 399

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
Db        400 CCCTTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGT 459

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205
Db        460 CGCTTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTC 519
```

```
Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGly 222
              |||||||||:::|||:::||||||:::|||:::|||           ::::::|||
Db        520 ATAAGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACT 579

Qy        223 LeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeu 242
              |||   :::   ::::::         |||||      :::||||||
Db        580 CTGGGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCA 630

Qy        243 TyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThr 262
              ::::::|||      :::   :::|||         :::         :::||||||
Db        631 GGGGTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACT 690

Qy        263 ThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPhe 282
              |||||||||||||||   |||           :::       |||||||||
Db        691 ACCTGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTC 747

Qy        283 IleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAla 302
              :::   :::         |||   :::|||   ::::::         ||||||
Db        748 TCAGCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTG 807

Qy        303 AlaIleIleArgThrLeuAsn---------------------AlaTyrAspHisArgTrp 315
              :::|||||||||   |||:::                     :::         |||
Db        808 TCTATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCT 867

Qy        316 LeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 335
              |||:::         |||       :::      ||||||   |||||||||   |||
Db        868 TTGTTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCC 918

Qy        336 SerAsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGly 354
              :::|||::::::||||||   |||::::::         :::   :::   |||:::
Db        919 ACAAACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCT 978

Qy        355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
              ||||||   ||||||:::   :::|||:::   |||::::::||||||:::||||||
Db        979 GCCTACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCC 1038

Qy        375 PheLeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1039 CTAATTTACTATTACGCTTCC 1059

RESULT 33
AAA35310
ID   AAA35310 standard; DNA; 3182 BP.
XX
AC   AAA35310;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human adenosine receptor related polynucleotide 2nd SEQ ID NO:184.
XX
KW   Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW   phosphorothioate; impaired respiration; inflammation; allergy;
KW   allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW   antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW   lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW   respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW   pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW   cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009525-A2.
XX
PD   24-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US17712.
XX
PR   03-AUG-1998;   98US-0095212.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
XX
PI   Nyce JW;
```

```
XX
DR   WPI; 2000-205971/18.
XX
PT   New antisense oligonucleotides useful for treating e.g. pulmonary
PT   vasoconstriction, inflammation, allergies, asthma, hypertension,
PT   bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT   cancers  -
XX
PS   Disclosure; Page 1330-1331; 1343pp; English.
XX
CC   The present invention describes a new composition comprising an
CC   antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC   targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC   inflammation. The ON can have antiinflammatory, antiallergic,
CC   antiasthmatic, cytostatic and analgesic activities. The compositions are
CC   useful for the treatment of diseases associated with inflammation,
CC   impaired airways, including lung disease and diseases whose secondary
CC   effects afflict the lungs of a subject. They can be used for treating
CC   e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC   asthma, impeded respiration, respiratory distress syndrome, pain, cystic
CC   fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC   pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC   carcinomas, and cancers which may metastasise to the lungs, including
CC   breast and prostate cancer. The reduction of the adenosine content of
CC   the ONs reduces side effects. The A-containing ONs break down with the
CC   release of deoxyadenosine which activates adenosine receptors causing
CC   bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC   nucleotide sequences given in the sequence listing from the present
CC   invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC   185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC   differ from the previously named sequences. SEQ ID NO:11 to 1680
CC   (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC   invention. N.B. Sequences given in the disclosure of the present
CC   invention do not match up with their corresponding SEQ ID NO: sequences
CC   given in the sequence listing.
XX
SQ   Sequence 3182 BP; 895 A; 701 C; 684 G; 902 T; 0 other;
```

```
Alignment Scores:
Pred. No.:                2.6e-37       Length:          3182
Score:                    512.00        Matches:         118
Percent Similarity:       54.47%        Conservative:    71
Best Local Similarity:    34.01%        Mismatches:      133
Query Match:              23.74%        Indels:          26
DB:                       21            Gaps:            9

US-09-208-629F-6 (1-408) x AAA35310 (1-3182)

Qy         52 ProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPhePro 71
              |||   :::|||          |||   |||   |||   :::   :::|||   |||
Db         44 CCTTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAAT-GATAAATATGAACCATTTTGG 102

Qy         72 PheSerAlaLeuGluGlyTrpThrGlyAlaThr------------IleThrValLysIle 87
                        :::   :::|||   |||             :::::::::
Db        103 GAGGATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAA 162

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                 |||   :::                    :::                |||||||||
Db        163 AGCAGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTG 222

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
              |||||||||       |||      :::|||::::::|||       ||||||||||||
Db        223 ACCAGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGC 282

Qy        128 ValProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
              :::|||   |||      |||:::   :::::::::|||      :::   :::
Db        283 CTCCCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG--- 339

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
                       |||:::      :::||||||   ||||||   ||||||         |||
Db        340 CCGGCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTC 399

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
```

```
Db        400 CCCTTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGT 459

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205

Db        460 CGCTTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTC 519

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGly 222

Db        520 ATAAGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACT 579

Qy        223 LeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeu 242

Db        580 CTGGGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCA 630

Qy        243 TyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThr 262

Db        631 GGGGTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACT 690

Qy        263 ThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPhe 282

Db        691 ACCTGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTC 747

Qy        283 IleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAla 302

Db        748 TCAGCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTG 807

Qy        303 AlaIleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrp 315

Db        808 TCTATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCT 867

Qy        316 LeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 335

Db        868 TTGTTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCC 918

Qy        336 SerAsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGly 354

Db        919 ACAAACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCT 978

Qy        355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374

Db        979 GCCTACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCC 1038

Qy        375 PheLeuTyrPheLeuMetSer 381

Db       1039 CTAATTTACTATTACGCTTCC 1059
```

RESULT 34
ABK70887
ID ABK70887 standard; cDNA; 1209 BP.
XX
AC ABK70887;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cDNA encoding PAR1 type thrombin receptor delta 1-49TR.
XX
KW Human; ss; gene; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.
XX
OS Homo sapiens.
XX
PN JP2002010784-A.
XX
PD 15-JAN-2002.
XX
PF 29-JUN-2000; 2000JP-0196514.
XX
PR 29-JUN-2000; 2000JP-0196514.
XX
PA (TEIJ) TEIJIN LTD.
XX
DR WPI; 2002-321520/36.
DR P-PSDB; ABG35298.
XX
PT An inhibitor of cell growth mediated by thrombin used to treat
PT inflammatory and cell proliferative diseases -
XX
PS Example 1; Page 22-24; 44pp; Japanese.
XX
CC The invention relates to a polypeptide or a compound which can inhibit
CC cell growth caused by thrombin. The polypeptide/compound combines to a
CC specific region of the structure of PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a cDNA encoding a PAR1 type thrombin receptor (or a
CC modified version).
XX
SQ Sequence 1209 BP; 254 A; 330 C; 273 G; 352 T; 0 other;

Alignment Scores:

Pred. No.:	1.94e-37	Length:	1209
Score:	507.50	Matches:	121
Percent Similarity:	52.32%	Conservative:	71
Best Local Similarity:	32.97%	Mismatches:	138
Query Match:	23.53%	Indels:	37
DB:	24	Gaps:	11

US-09-208-629F-6 (1-408) x ABK70887 (1-1209)

```
Qy         32 LeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLys 51

Db         16 CTGCTGCTGGTGGCCGCCTGCTTCAGT----------------------CTGTGCGGC 51

Qy         52 ProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPhePro 71

Db         52 CCGCTGTTGTCTGCCCGCACC-----------CGGGCCGATAAATATGAACCATTTTGG 99

Qy         72 PheSerAlaLeuGluGlyTrpThrGlyAlaThr-----------IleThrValLysIle 87

Db        100 GAGGATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAA 159

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107

Db        160 AGCAGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTG 219

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127

Db        220 ACCAGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGC 279

Qy        128 ValProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145

Db        280 CTCCCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG--- 336

Qy        146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165

Db        337 CCGGCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTC 396

Qy        166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185

Db        397 CCCTTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGT 456

Qy        186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205

Db        457 CGCTTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTC 516

Qy        206 IleSerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGly 222

Db        517 ATAAGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACT 576
```

```
Qy        223 LeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeu 242
              |||   :::   ::::::         ||||||      :::|||||||
Db        577 CTGGGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCA 627

Qy        243 TyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThr 262
              ::::::|||      :::   :::|||         :::         :::|||||||
Db        628 GGGGTAGTGCCTCTGCTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACT 687

Qy        263 ThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPhe 282
              ||||||||||||||   |||                  :::      |||||||||
Db        688 ACCTGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTC 744

Qy        283 IleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAla 302
              :::   :::         |||   :::|||   ::::::         ||||||
Db        745 TCAGCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCACGGTCTGTTATGTG 804

Qy        303 AlaIleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrp 315
              :::|||||||||   |||:::                  :::       |||
Db        805 TCTATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCT 864

Qy        316 LeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 335
              |||:::          |||      :::      ||||||   ||||||||   |||
Db        865 TTGTTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGACCC 915

Qy        336 SerAsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGly 354
              :::|||::::::||||||   |||:::::         :::   :::   |||:::
Db        916 ACAAACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCT 975

Qy        355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
              ||||||   ||||||:::   :::|||:::   |||::::::   |||:::||||||
Db        976 GCCTACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTGCTGCATCGACCCC 1035

Qy        375 PheLeuTyrPheLeuMetSer 381
              :::|||:::      |||
Db       1036 CTAATTTACTATTACGCTTCC 1056
```

```
RESULT 35
AAQ28568
ID   AAQ28568 standard; DNA; 3480 BP.
XX
AC   AAQ28568;
XX
DT   15-FEB-1993  (first entry)
XX
DE   Human thrombin receptor gene.
XX
KW   Diagnosis; cardiovascular disease; wound healing; restenosis;
KW   thrombosis; unstable angina treatment; myocardial infarction;
KW   thrombotic; thromboembolytic stroke; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1502
FT                   /*tag=  a
FT   sig_peptide     225..303
FT                   /*tag=  b
FT   mat_peptide     304..1502
FT                   /*tag=  c
XX
PN   WO9214750-A.
XX
PD   03-SEP-1992.
XX
PF   19-FEB-1992;   92WO-US01312.
XX
PR   19-FEB-1991;   91US-0657769.
PR   07-NOV-1991;   91US-0789184.
XX
PA   (CORT-) COR THERAPEUTICS INC.
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Coughlin SR,  Scarborough RM;
XX
DR   WPI; 1992-316119/38.
DR   P-PSDB; AAR27240.
XX
PT   DNA encoding cell surface receptor for thrombin - useful for
PT   determining thrombin in diagnosing e.g. cardiovascular diseases,
PT   also to treat wound healing, restenosis etc.
XX
PS   Disclosure; Fig 1; 81pp; English.
XX
CC   The sequence is that of the gene coding for human thrombin receptor.
CC   It can be used in the prepn. of diagnostics to determine thrombin
CC   levels in samples, and screening tools for candidate substances which
CC   affect thrombin activity in vivo. Thrombosis may be diagnosed in a
CC   mammal by measuring the presence, absence or amt. of the cleaved
CC   activation peptide of the TR.
XX
SQ   Sequence 3480 BP; 947 A; 816 C; 786 G; 931 T; 0 other;
```

```
Alignment Scores:
Pred. No.:                1.6e-36       Length:          3480
Score:                    504.00        Matches:         125
Percent Similarity:       52.07%        Conservative:    76
Best Local Similarity:    32.38%        Mismatches:      155
Query Match:              23.37%        Indels:          30
DB:                       13            Gaps:            11

US-09-208-629F-6 (1-408) x AAQ28568 (1-3480)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||   :::   :::   :::   |||     |||:::     |||   ||||||
Db        228 GGGCCGCGGCGGCTGCTGCTGGTGGCCGCCTGCTTCAGTCTGTGCGGCCCGCTGTTGTCT 287

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||      :::                |||||                 |||
Db        288 GCCCGCACCCGGGCCCGCAGGCCAGAATCAAAAGCAACAAAT-----------GCCACC 335

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
              |||      ::::::|||             ||||||      :::|||   |||
Db        336 TTAGATCCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395

Qy         73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----------IleThrValLysIleLys 88
                 :::      :::|||   |||             :::::::::
Db        396 GATGAGGAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455

Qy         89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
              |||   :::                 :::                 ||||||||||
Db        456 AGTCCTCTTCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTCCGGATATTTGACC 515

Qy        109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
              ||||||      |||      :::|||:::::::|||      ||||||||||||   :::
Db        516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTTTGTAGTCAGCCTC 575

Qy        129 ProAlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
              |||   |||      |||:::   :::::::::|||      :::   :::
Db        576 CCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAAGGTCAAGAAG---CCG 632

Qy        147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
                 |||:::      :::||||||   ||||||   ||||||          ||||||
Db        633 GCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTTTGTGTCTGTGCTCCCC 692

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              |||||||||:::|||:::   :::|||:::::::|||   ||||||      |||||||||
Db        693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTCTGAATTGTGTCGC 752

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                 |||      ||||||   |||||||||   ||||||||||||:::      |||
Db        753 TTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTTGCTCATGACAGTCATA 812

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeu 223
              ||||||:::|||:::||||||:::|||:::|||             ::::::|||   |||
Db        813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCCTGGCGTACTCTG 872
```

```
Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
Db        873 GGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
Db        924 GTAGTGCCTCTCGTCCTCAAGGAGCAAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
Db        984 TGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTATGCCTACTACTTCTCA 1040

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
Db       1041 GCCTTCTCTGCTGTCTTCAAAAAAGTGCCGCTGATCATTTCCACGGTCTGTTATGTGTCT 1100

Qy        304 IleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrpLeu 316
Db       1101 ATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG 1160

Qy        317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
Db       1161 TTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCATTTGCTTCGGAGCCACA 1211

Qy        337 AsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
Db       1212 AACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTCCACCACAGAGGCTGCC 1271

Qy        356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
Db       1272 TACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCCTA 1331

Qy        376 LeuTyrPheLeuMetSer 381
Db       1332 ATTTACTATTACGCTTCC 1349
```

```
RESULT 36
ABK70888
ID   ABK70888 standard; cDNA; 1116 BP.
XX
AC   ABK70888;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human cDNA encoding PAR1 type thrombin receptor delta 1-80.
XX
KW   Human; ss; gene; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW   inflammatory disease; cell proliferative disease.
XX
OS   Homo sapiens.
XX
PN   JP2002010784-A.
XX
PD   15-JAN-2002.
XX
PF   29-JUN-2000; 2000JP-0196514.
XX
PR   29-JUN-2000; 2000JP-0196514.
XX
PA   (TEIJ ) TEIJIN LTD.
XX
DR   WPI; 2002-321520/36.
DR   P-PSDB; ABG35299.
XX
PT   An inhibitor of cell growth mediated by thrombin used to treat
PT   inflammatory and cell proliferative diseases -
XX
PS   Example 1; Page 24-25; 44pp; Japanese.
XX
CC   The invention relates to a polypeptide or a compound which can inhibit
CC   cell growth caused by thrombin. The polypeptide/compound combines to a
CC   specific region of the structure of PAR1 type human thrombin receptor
CC   participating to cell growth. Preferably, the compound contains the
CC   52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC   human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC   amino acid or peptide sequence). Also included are a modified PAR1 type
CC   thrombin receptor gene or its fragment used for obtaining the above
CC   polypeptide, a human PAR1 type thrombin receptor protein and its
CC   encoding DNA  comprising a fully. The polypeptide or the compound is used
CC   to treat inflammatory diseases and cell proliferative diseases. The
CC   present sequence is a cDNA encoding a PAR1 type thrombin receptor (or a
CC   modified version).
XX
SQ   Sequence 1116 BP; 218 A; 318 C; 252 G; 328 T; 0 other;
```

```
Alignment Scores:
Pred. No.:              4.09e-37      Length:          1116
Score:                  503.50        Matches:         112
Percent Similarity:     58.01%        Conservative:    69
Best Local Similarity:  35.90%        Mismatches:      105
Query Match:            23.34%        Indels:          26
DB:                     24            Gaps:            10

US-09-208-629F-6 (1-408) x ABK70888 (1-1116)

Qy         87 IleLysCysProGlu---------GluSerAlaSerHisLeuHisValLys-AsnAlaTh 103
Db         57 GTTGTCTGCCCGCACCCGGGCCCAAAAACAACTTCCTGCATTCATCTCAGAAGATGCCTC 116

Qy        103 rMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVa 123
Db        117 C---GGATATTTGACCAGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGT 173

Qy        123 lPheValValGlyValProAlaAsn------AlaValThrLeuTrpMetLeuPhePheAr 141
Db        174 GTTTGTAGTCAGCCTCCCACTAAACATCATGGCCATCGTTGTGTTCATCCTGAAAATGAA 233

Qy        141 gThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPh 161
Db        234 GGTCAAGAAG---CCGGCGGTGGTGTACATGCTGCACCTGGCCACGGCAGATGTGCTGTT 290

Qy        161 eCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGl 181
Db        291 TGTGTCTGTGCTCCCCTTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGG 350

Qy        181 yGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLe 201
Db        351 GTCTGAATTGTGTCGCTTCGTCACTGCAGCATTTTACTGTAACATGTACGCCTCTATCTT 410

Qy        201 uLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisPro---------Ph 218
Db        411 GCTCATGACAGTCATAAGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCT 470

Qy        218 eThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAl 238
Db        471 CTCCTGGCGTACTCTGGGAAGGGCTTCCTTC---------ACTTGTCTGGCCATCTGGGC 521

Qy        238 aThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGl 258
Db        522 TTTGGCCATCGCAGGGGTAGTGCCTCTGCTCCTCAAGGAGCAAACCATCCAGGTGCCCGG 581

Qy        258 nProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGl 278
Db        582 GCTCAACATCACTACCTGTCATGATGTGCTCAATGAAACCCTGCTCGAAGGC---TACTA 638

Qy        278 nLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIl 298
Db        639 TGCCTACTACTTCTCAGCCTTCTCTGCTGTCTTCTTTTTTGTGCCGCTGATCATTTCCAC 698

Qy        298 eTyrCysTyrAlaAlaIleIleArgThrLeuAsn--------------------AlaTy 311
Db        699 GGTCTGTTATGTGTCTATCATTCGATGTCTTAGCTCTTCCGCAGTTGCCAACCGCAGCAA 758

Qy        311 rAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIl 331
Db        759 GAAGTCCCGGGCTTTGTTC---------CTGTCAGCTGCTGTTTTCTGCATCTTCATCAT 809
```

```
Qy        331 eCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyr---TyrTyrAs 350
Db        810 TTGCTTCGGACCCACAAACGTCCTCCTGATTGCGCATTACTCATTCCTTTCTCACACTTC 869

Qy        350 nAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSe 370
Db        870 CACCACAGAGGCTGCCTACTTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAAGCTG 929

Qy        370 rCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db        930 CTGCATCGACCCCCTAATTTACTATTACGCTTCC 963
```

RESULT 37
AAT93367
ID AAT93367 standard; cDNA; 1312 BP.
XX
AC AAT93367;
XX
DT 01-MAY-1998 (first entry)
XX
DE Mouse thrombin receptor cDNA.
XX
KW Transgenic animal; thrombin receptor; animal model; restenosis;
KW knockout animal; antagonist; agonist; mouse; murine;
KW Alzheimer's disease; wound repair; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT primer_bind complement (612..631)
FT /*tag= a
FT /label= MTR-3
FT primer_bind 1294..1311
FT /*tag= b
FT /label= MTR-2
XX
PN WO9740138-A1.
XX
PD 30-OCT-1997.
XX
PF 22-APR-1997; 97WO-US06575.
XX
PR 25-JUN-1996; 96US-0020544.
PR 23-APR-1996; 96US-0016070.
XX
PA (ORTH) ORTHO PHARM CORP.
XX
PI Andrade-Gordon P, Leung W, Zhou L;
XX
DR WPI; 1997-535823/49.
XX
PT Transgenic animals with altered form of thrombin receptor gene -
PT useful in study of receptor function and identification of
PT antagonists and agonists for disease therapy
XX
PS Example 1; Fig 1; 47pp; English.
XX
CC The present sequence was used in the preparation of a transgenic
CC non-human vertebrate with somatic and germ cells containing an
CC altered form of a thrombin receptor (TR) gene, where the altered
CC gene has replaced a wild type TR gene in the animal or an ancestor
CC at an embryonic stage using embryonic stem cells. The transgenic
CC animal provides an animal model to understand receptor function in,
CC e.g. platelets, endothelial cells and neurons, and evaluate drug
CC therapies modulating TR function/expression in human cells.
CC Alterations rendering the wild type gene non-functional produce
CC knockout animals, useful as non-human whole animal models for
CC diseases involving TR equivalents in humans. Alternatively,
CC substituting the naturally occurring gene for a gene from a 2nd
CC species or with a mutation, results in animals producing the 2nd
CC species or mutated gene products. These transgenic animals are
CC useful for drug antagonist and agonist studies, creation of animal
CC models of human diseases and investigation of disorders associated
CC with human TR-mediated responses. Antagonists may be used
CC therapeutically for cardiovascular problems, e.g. restenosis, to
CC limit platelet adhesion during interventional procedures or to
CC generate therapies for bone degenerative and neurodegenerative
CC diseases, e.g. Alzheimer's. Agonists may be useful in wound repair.
XX
SQ Sequence 1312 BP; 247 A; 406 C; 328 G; 331 T; 0 other;

Alignment Scores:

Pred. No.:	2.23e-36	Length:	1312
Score:	496.50	Matches:	127
Percent Similarity:	49.53%	Conservative:	82
Best Local Similarity:	30.09%	Mismatches:	164
Query Match:	23.02%	Indels:	49
DB:	18	Gaps:	16

US-09-208-629F-6 (1-408) x AAT93367 (1-1312)

```
Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPhe---CysGln 39
Db         24 CGGCGCTTGCTGATCGTCGCCCTCGGCCTCAGCCTGTGCGGTCCCTTGCTGTCTTCCCGC 83

Qy         40 SerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db         84 GTCCCTATGAGCCAGCCAGAATCAGAGAGGACAGATGCTACGGTGAACCCCCGCTCATTC 143

Qy         60 ------ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGly 77
Db        144 TTTCTAAGGAATCCCAGTGAAAATACATTTGAACTGGTCCCCCTGGGGGATGAGGAGGAG 203

Qy         78 ----------------------TrpThrGlyAlaThrIleThrValLysIleLysCys 89
Db        204 GAGGAGAAAAATGAAAGCGTCCTGCTGGAGGGTAGGGCAGTCTACTTAAATATAAGCCTC 263

Qy         90 ProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSer 109
Db        264 CCTCCTCACACGCCGCCTCCTCCCTTCATCTCCGAGGACGCCTCCGGATATCTGACCAGC 323

Qy        110 SerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValPro 129
Db        324 CCCTGGCTGACGCTCTTCATGCCCTCCGTGTACACGATTGTGTTCATTGTCAGCCTTCCT 383

Qy        130 AlaAsn------AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
Db        384 CTGAACGTCCTGGCCATCGCAGTGTTCGTCTTGAGGATGAAGGTCAAGAAG---CCGGCC 440

Qy        148 ThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPhe 167
Db        441 GTGGTGTACATGCTGCACCTGGCCATGGCCGACGTGCTCTTCGTGTCGGTGCTCCCTTCC 500

Qy        168 LysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAla 187
Db        501 AAGATCAGCTACTACTTCTCCGGCACTGATTGGCAGTTCGGGTCTGGAATGTGCCGTTTC 560

Qy        188 ThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSer 207
Db        561 GCCACCGCAGCGTTTTACTATAACATGTACGCCTCCATCATGCTCATGACGGTCATAAGC 620

Qy        208 IleAsnArgTyrLeuAlaIleValHisPro---------PheThrTyrArgGlyLeuPro 224
Db        621 ATTGACCGGTTCCTGGCGGTGGTGTATCCGATCCAGTCCCTGTCCTGGCGCACTCTGGGC 680

Qy        225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
Db        681 GGAGCCAACTTC---------ACTTGCGTGGTCATTTGGGTGATGGCCATCATGGGGGTG 731

Qy        245 LeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCys 264
Db        732 GTGCCCCTTCTCCTCAAGGAGCAGACCACCCGAGTTCCGGGACTCAACATCACCACCTGC 791

Qy        265 HisAspVal------HisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPhe 282
```

```
Db        792 CACGATCTCCTCAGTGAGAACCTGATGCAAGGC---------TTTTACTCGTACTACTTC 842

Qy        283 IleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAla 302
Db        843 TCGGCCTTCTCCGCCATCTTCTTTCTTGTGCCGTTGATCGTTTCCACGGTCTGCTACACG 902

Qy        303 AlaIleIleArgThrLeuAsn--------------------AlaTyrAspHisArgTrp 315
Db        903 TCCATCATCCGGTGCCTGAGCTCCTCCGCGGTTGCCAACCGGAGCAAGAAGTCGCGGGCT 962

Qy        316 LeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 335
Db        963 TTGTTC---------CTGTCTGCCGCGGTGTTCTGCATCTTCATCGTCTGCTTTGGGCCC 1013

Qy        336 SerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyr------------AsnAsn 351
Db       1014 ACCAACGTCCTCCTGATTGTGCAC---------TACCTTTTCCTCTCCGACAGTCCTGGT 1064

Qy        352 ThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db       1065 ACGGAGGCAGCCTACTTTGCTTACCTCCTCTGCGTCTGTGTGACGAGCGTGAGCTGCTGC 1124

Qy        372 LeuAspProPheLeu---TyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyr 390
Db       1125 ATCGATCCGTTGATTTACTACTACGCCTCCTCCGAGTGCCAGAGGCACCTCTACAGCATC 1184

Qy        391 Leu------------ThrLys******AsnAspLeuArgGluGlnGlyGln------Pro 404
Db       1185 TTGTGCTGCAAAGAAAGCTCTGATCCCAACAGTTGCAACAGCACCGGCCAGCTGATGCCG 1244

Qy        405 SerGln 406
Db       1245 AGTAAA 1250
```

```
RESULT 38
AAV40372
ID   AAV40372 standard; cDNA; 2588 BP.
XX
AC   AAV40372;
XX
DT   16-OCT-1998  (first entry)
XX
DE   Mouse G-protein coupled receptor 69A08 #1 encoding cDNA.
XX
KW   Mouse; CKDLR201.1; chemokine; G-protein coupled receptor; GPCR;
KW   7 transmembrane receptor; inflammation; asthma; antiviral;
KW   abnormal cell proliferation; regeneration; degeneration; atrophy; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1083
FT                   /*tag=  a
FT                   /product= "G-protein coupled receptor 69A08"
FT   misc_feature    158
FT                   /*tag=  b
FT                   /note= "residues 158, 159 and 276 are probably absent,
FT                          changing the reading frame between those
FT                          positions; sequence are provided in AAV40373 and
FT                          AAW69598"
XX
PN   WO9831810-A2.
XX
PD   23-JUL-1998.
XX
PF   20-JAN-1998;   98WO-US00218.
XX
PR   21-JAN-1997;   97US-0786624.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Capone M,  Gorman DM,  Hedrick JA,  Huffine CF,  Rossi DL;
PI   Vicari A,  Zlotnik A;
XX
DR   WPI; 1998-414108/35.
DR   P-PSDB; AAW69597.
XX
PT   Pure or recombinant chemokine CKDLR201.1 - useful, e.g. for treating
PT   inflammation and as antiviral agents
XX
PS   Claim 7; Page 61-64; 77pp; English.
XX
CC   The present sequence encodes mouse G-protein coupled receptor 69A08 #1
CC   protein which comprises a plurality of epitopes. Host cells containing
CC   vectors comprising a nucleotide sequence encoding the protein are used to
CC   produce recombinant protein. Treatment of a cell, particularly neuron,
CC   macrophage or lymphocyte, carrying a G-protein coupled receptor (GPCR)
CC   or a receptor responsive to CKDLR201.1 chemokine protein, with an
CC   (ant)agonist is used to control physiological development, e.g.
CC   alteration of calcium ion influx, a chemoattractant response, morphology,
CC   phosphoinositide lipid turnover or an antiviral response. Nucleotide
CC   sequences the CKDLR201.1 protein, or GPCRs, are useful as primers or
CC   probes, e.g. for detecting and isolating related sequences and for
CC   expressing antigenic peptides. Antibodies (Ab) directed against the
CC   CKDLR201.1 protein and GPCRs are used to detect or purify the proteins;
CC   diagnostically (e.g. for developmental abnormalities); in screening for
CC   potential drugs; to inhibit chemokine/receptor activation; (when coupled
CC   to a toxin or radioisotope) for killing specific cells, and to raise
CC   anti-idiotype antibodies. CKDLR201.1 protein and GPCRs and compounds
CC   which bind them can be used to treat inflammation, e.g. asthma; as
CC   antiviral agents, and to treat abnormal cell proliferation, regeneration,
CC   degeneration and atrophy. Therapeutic agents are administered orally, by
CC   injection and rectally.
XX
SQ   Sequence 2588 BP; 462 A; 782 C; 743 G; 601 T; 0 other;
```

Alignment Scores:			
Pred. No.:	3.28e-35	Length:	2588
Score:	488.00	Matches:	117
Percent Similarity:	51.45%	Conservative:	60
Best Local Similarity:	34.01%	Mismatches:	134
Query Match:	22.62%	Indels:	34
DB:	19	Gaps:	4

US-09-208-629F-6 (1-408) x AAV40372 (1-2588)

```
Qy         56 IleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeu 75
Db       1135 GTAAAGAGTACCAGGGGAAGCCAT---------------------------------- 1158

Qy         76 GluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer------ 93
Db       1159 GAAGGCCCTCTGGGTCCC------ACAGTACAACTCAAGGAGCCGAAGTCCTCAGACAAG 1212

Qy         94 ----------------------------------------AlaSerHisLeuHisVal 99
Db       1213 CCTAATCCACGAGGCTACCCGGGCAAATTCTGTGCCAACGACAGTGACACGCTGGAGCTC 1272

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
Db       1273 CCGGCCAGCTCTCAAGCACTGCTGCTGGGGTGGGTCCCCACGAAGCTGGTACCTGCCCTC 1332

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db       1333 TATGGGCTTGTGGTGGCTGTGGGGCTGCCTGCCAATGGGCTGGCGCTGTGGGTGCTGGCC 1392

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyrThr-AsnLeuAlaIleAlaAspPh 159
Db       1393 ACAAGGGTGCCACGCCTGCCATCCACCATTCTGCTCATGGAACCTGGCAGTGGCTGATCT 1452

Qy        159 eLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVa 179
Db       1453 GCTGTTGGCCCTGGTGCTGCCACCACGACTGGCTTACCACTTGCGTGGCCAGCGCTGGCC 1512

Qy        179 lPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSe 199
```

```
Db       1513 ATTTGGTGAGGCTGCCTGCCGGGTGGCCACAGCTGCCCTCTATGGCCACATGTATGGTTC 1572

Qy        199 rIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheTh 219
              |:::|||||||||||||   :::|||:::::::|||||||||||:::|||||||||
Db       1573 AGTGTTGCTGCTGGCTGCAGTCAGCTTGGACAGATACCTGGCCCTGGTGCATCCTTTGCG 1632

Qy        219 rTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaTh 239
                  |||   |||                                 |||   :::   |||   ::
Db       1633 GGCCCGTGCACTGCGTGGTCAACGCCTCACTACTGGACTCTGTTTGGTGGCCTGGCTCTC 1692

Qy        239 rValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPr 259
              :                 ||||||      :::   :::|||   :::   |||:::   ||
Db       1693 TGCAGCCACCCTGGCCTTGCCTCTCACTCTGCATCGGCAGAACTTCCGATTACTGGCTCC 1752

Qy        259 oAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLe 279
              |           ||||||||                 |||   :::|||   ::::::
Db       1753 GAT-CGCATGCTGTGTCATGATGCGCTGCCCCTGGCTGAGCAGAACTCCCACTGGAGAAC 1811

Qy        279 uTyrTyrPheIleSerLeuAlaPhePheGlyPheLeu-IleProPheValLeuIleIleT 299
                  ||||||   ||||||       |||       :::|||   :::   :::
Db       1812 G---GCCTTCATCTGCCTGGCTGTCCTGGGCTGCTTCCTTGCCACTGCTGGCCATGGGCC 1868

Qy        299 yrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrV 319
              ||||||        :::|||   |||   |||                ||:::      :
Db       1869 TGTGCTATGGAACCACCCTTCGTGCATTGGCGGCCAATGGCCAGCGCTACAGCCATGCAC 1928

Qy        319 alLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleI 339
              :::::   :::   |||:::|||                  |||   |||||||||:::
Db       1929 TCAGACTGACAGCCCTGGTACTGTTCTCGGCAGTGGCTTCTTTCACACCTAGCAATGTGC 1988

Qy        339 leLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleT 359
              ::|||::::::|||::::::|||                          ||||||      |
Db       1989 TGCTGGTGCTGCACTATTCAAACCCGAGCCCTGAGGCCTGGGGCAATCTCTATGGAGCCT 2048

Qy        359 yrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheL 379
              ||:::   :::|||   |||   :::||||||||||||:::||||||||:::|||:::
Db       2049 ATGTGCCCAGCCTGGCACTCAGCACCCTCAACAGCTGCGTAGACCCTTTCATCTACTACT 2108

Qy        379 euMetSer 381
                :::|||
Db       2109 ATGTGTCC 2116
```

```
RESULT 39
AAV40373
ID   AAV40373 standard; cDNA; 1080 BP.
XX
AC   AAV40373;
XX
DT   16-OCT-1998  (first entry)
XX
DE   Mouse G-protein coupled receptor 69A08 #2 encoding cDNA.
XX
KW   Mouse; CKDLR201.1; chemokine; G-protein coupled receptor; GPCR;
KW   7 transmembrane receptor; inflammation; asthma; antiviral;
KW   abnormal cell proliferation; regeneration; degeneration; atrophy; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1080
FT                   /*tag=  a
FT                   /product= "G-protein coupled receptor 69A08"
XX
PN   WO9831810-A2.
XX
PD   23-JUL-1998.
XX
PF   20-JAN-1998;   98WO-US00218.
XX
PR   21-JAN-1997;   97US-0786624.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Capone M,  Gorman DM,  Hedrick JA,  Huffine CF,  Rossi DL;
PI   Vicari A,  Zlotnik A;
XX
DR   WPI; 1998-414108/35.
DR   P-PSDB; AAW69598.
XX
PT   Pure or recombinant chemokine CKDLR201.1 - useful, e.g. for treating
PT   inflammation and as antiviral agents
XX
PS   Claim 7; Page 66-68; 77pp; English.
XX
CC   The present sequence encodes mouse G-protein coupled receptor 69A08 #2
CC   protein which comprises a plurality of epitopes. Host cells containing
CC   vectors comprising a nucleotide sequence encoding the protein are used to
CC   produce recombinant protein. Treatment of a cell, particularly neuron,
CC   macrophage or lymphocyte, carrying a G-protein coupled receptor (GPCR)
CC   or a receptor responsive to CKDLR201.1 chemokine protein, with an
CC   (ant)agonist is used to control physiological development, e.g.
CC   alteration of calcium ion influx, a chemoattractant response, morphology,
CC   phosphoinositide lipid turnover or an antiviral response. Nucleotide
CC   sequences the CKDLR201.1 protein, or GPCRs, are useful as primers or
CC   probes, e.g. for detecting and isolating related sequences and for
CC   expressing antigenic peptides. Antibodies (Ab) directed against the
CC   CKDLR201.1 protein and GPCRs are used to detect or purify the proteins;
CC   diagnostically (e.g. for developmental abnormalities); in screening for
CC   potential drugs; to inhibit chemokine/receptor activation; (when coupled
CC   to a toxin or radioisotope) for killing specific cells, and to raise
CC   anti-idiotype antibodies. CKDLR201.1 protein and GPCRs and compounds
CC   which bind them can be used to treat inflammation, e.g. asthma; as
CC   antiviral agents, and to treat abnormal cell proliferation, regeneration,
CC   degeneration and atrophy. Therapeutic agents are administered orally, by
CC   injection and rectally.
XX
SQ   Sequence 1080 BP; 177 A; 347 C; 307 G; 249 T; 0 other;
```

Alignment Scores:

Pred. No.:	1.46e-35	Length:	1080
Score:	486.50	Matches:	111
Percent Similarity:	52.96%	Conservative:	59
Best Local Similarity:	34.58%	Mismatches:	135
Query Match:	22.55%	Indels:	17
DB:	19	Gaps:	2

US-09-208-629F-6 (1-408) x AAV40373 (1-1080)

```
Qy         76 GluGlyTrpThrGlyAlaThrIleThrValLys-IleLysCysProGluGluSerAlaSe 95
              ||||||      |||   |||:::   :::|||    :::   |||:::   |||
Db          3 GAAGGCCCTCTGGGTCCCACAGTACAACTCAAGGAGCCGAAGTCCTCAGACAAGCCTAAT 62

Qy         95 rHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLys-- 114
              ·|||         :::      ::::::              :::      |||::::::
Db         63 CCACGAGGCTACCCGGGCAAATTCTGTGCCAACGACAGTGACACGCTGGAGCTCCCGGCC 122

Qy        115 ------------------------------------LeuIleProAlaIleTyrLeuLe 122
                                                  |||:::||||||:::|||   ||
Db        123 AGCTCTCAAGCACTGCTGCTGGGGTGGGTCCCCACAGCTGGTACCTGCCCTCTATGGGCT 182

Qy        122 uValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgTh 142
              ||||      ||||||:::|||||||||   :::   ||||||:::|||      |||
Db        183 TGTGGTGGCTGTGGGGCTGCCTGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACAAGGGT 242

Qy        142 rArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCy 162
                  :::   :::|||:::         ||||||||:::||||||   |||
Db        243 GCCACGCCTGCCATCCACCATTCTGCTCATGAACCTGGCAGTGGCTGATCTGCTGTTGGC 302

Qy        162 sValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGl 182
               :::   ||||||   ::::::||||||||||||   |||      |||   ||||||||
Db        303 CCTGGTGCTGCCACCACGACTGGCTTACCACTTGCGTGGCCAGCGCTGGCCATTTGGTGA 362

Qy        182 uValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLe 202
              |      ||||||      |||          ||||||:::||||||   |||:::||||||
Db        363 GGCTGCCTGCCGGGTGGCCACAGCTGCCCTCTATGGCCACATGTATGGTTCAGTGTTGCT 422
```

```
Qy        202 uLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGl 222
Db        423 GCTGGCTGCAGTCAGCTTGGACAGATACCTGGCCCTGGTGCATCCTTTGCGGGCCCGTGC 482

Qy        222 yLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLe 242
Db        483 ACTGCGTGGTCAACGCCTCACTACTGGACTCTGTTTGGTGGCCTGGCTCTCTGCAGCCAC 542

Qy        242 uTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleTh 262
Db        543 CCTGGCCTTGCCTCTCACTCTGCATCGGCAGAACTTCCGATTACTGGCTCCGAT-CGCAT 601

Qy        262 rThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPh 282
Db        602 GCTGTGTCATGATGCGCTGCCCCTGGCTGAGCAGAACTCCCACTGGAGAACG---GCCTT 658

Qy        282 eIleSerLeuAlaPhePheGlyPheLeu-IleProPheValLeuIleIleTyrCysTyrA 302
Db        659 CATCTGCCTGGCTGTCCTGGGCTGCTTCCTTGCCACTGCTGGCCATGGGCCTGTGCTATG 718

Qy        302 laAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaS 322
Db        719 GAACCACCCTTCGTGCATTGGCGGCCAATGGCCAGCGCTACAGCCATGCACTCAGACTGA 778

Qy        322 erLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleI 342
Db        779 CAGCCCTGGTACTGTTCTCGGCAGTGGCTTCTTTCACACCTAGCAATGTGCTGCTGGTGC 838

Qy        342 leHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleA 362
Db        839 TGCACTATTCAAACCCGAGCCCTGAGGCCTGGGGCAATCTCTATGGAGCCTATGTGCCCA 898

Qy        362 laLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db        899 GCCTGGCACTCAGCACCCTCAACAGCTGCGTAGACCCTTTCATCTACTACTATGTGTCC 957
```

RESULT 40
AAZ99588
ID AAZ99588 standard; DNA; 1955 BP.
XX
AC AAZ99588;
XX
DT 03-JUL-2000 (first entry)
XX
DE cDNA encoding a human 14400 receptor polypeptide.
XX
KW Human; 14400 receptor; G-protein coupled receptor; signalling pathway;
KW CD8 T cell; CD4 T cell; CD34+ bone marrow cell; thrombocytopenia;
KW inflammation; spleen disorder; splenomegaly; lung disorder;
KW adult respiratory distress syndrome; colon disorder;
KW bacterial enterocolitis; liver disorder; hepatic injury;
KW platelet number; precursor T-cell neoplasm; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 288..1367
FT /*tag= a
FT /product= "14400 receptor"
FT /transl_except= (pos: 1119, 1121, aa: Thr)
XX
PN WO200011170-A1.
XX
PD 02-MAR-2000.
XX
PF 20-AUG-1999; 99WO-US19112.
XX
PR 20-AUG-1998; 98US-0137063.
PR 20-AUG-1999; 99US-0378100.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Glucksmann MA, Weich NS;
XX
DR WPI; 2000-256378/22.
DR P-PSDB; AAY69485.
XX
PT G-protein coupled receptor, 14400 receptor, for treating CD4, CD8 or
PT CD34+ related disorders -
XX
PS Claim 3; Fig 1A-B; 92pp; English.
XX
CC The present sequence encodes a human 14400 receptor polypeptide. The
CC receptor is a G-protein coupled receptor that participates in signalling
CC pathways. The polypeptide can be used to identify an agent to modulate
CC the polypeptides activity in patients having a disorder involving CD8 or
CC CD4 T cells or CD34+ bone marrow cells, especially where the disorder
CC involves thrombocytopenia or inflammation. The modulators can
CC be used to treat disorders involving the spleen (e.g. splenomegaly),
CC lung (e.g. adult respiratory distress syndrome), colon (e.g. bacterial
CC enterocolitis), liver (e.g. hepatic injury), the uterus and
CC endometrium, brain, T-cells, skin, heart, blood vessels, red cells,
CC thymus, B-cells, kidney, breast, testis and epididymis, prostate,
CC thyroid, skeletal muscle, pancreas, small intestine, reduced platelet
CC number, precursor T-cell neoplasms, CD3, CD4, and CD8 T lymphocytes.
XX
SQ Sequence 1955 BP; 322 A; 642 C; 567 G; 421 T; 3 other;

Alignment Scores:

Pred. No.:	2.46e-33	Length:	1955
Score:	466.00	Matches:	104
Percent Similarity:	51.95%	Conservative:	56
Best Local Similarity:	33.77%	Mismatches:	130
Query Match:	21.60%	Indels:	18
DB:	21	Gaps:	8

US-09-208-629F-6 (1-408) x AAZ99588 (1-1955)

```
Qy         90 ProGluGluSerAlaSerHisLeuHisVal----------LysAsnAlaThr 103
Db        267 CCGGGCGAGGACCCCTCCAGGATGCAGGTCCCGAACAGCACCGGCCCGGACAACGCGACG 326

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
Db        327 CTGCAGATGCTGCGGAACCCGGCGATCGCGGTGGCCCTGCCCGTGGTGTACTCGCTGGTG 386

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg------ 141
Db        387 GCGGCGGTCAGCATCCCGGGCAACCTCTTCTCTCTGTGGGTGCTGTGCCGGCGCATGGGG 446

Qy        142 ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPhe 161
Db        447 CCCAGATCC---CCGTCGGTCATCTTCATGATCAACCTGAGCGTCACGGACCTGATGCTG 503

Qy        162 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 181
Db        504 GCCAGCGTGTTGCCTTTCCAAATCTACTACCATTGCAACCGCCACCACTGGGTATTCGGG 563

Qy        182 GluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeu 201
Db        564 GTGCTGCTTTGCAACGTGGTGACCGTGGCCTTTTACGCAAACATGTATTCCAGCATCCTC 623

Qy        202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
Db        624 ACCATGACCTGTATCAGCGTGGAGCGCTTCCTGGGGGTCCTGTACCCGCTCAGCTCCAAG 683

Qy        222 GlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
Db        684 CGCTGGCGCCGCCGTCGTTACGCGGTGGCCGCGTGTGCAGGGACCTGGCTGCTGCTCCTG 743

Qy        242 LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db        744 ACCGCCCTGTCCCCGCTGGCGCGCACCGATCTCACCTACCCGGTGCACGCCCTGGGCATC 803

Qy        262 ThrThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPheGlnLeuTyr 280
```

```
Db        804 ATCACCTGCTTCGACGTCCTCAAGTGGACGATGCTCCCCAGCGTGGCCATGTGGGCCGTG 863

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              :::      ::::::           ||||||||||||||||||:::   :::   |||
Db        864 TTCCTCTTCACCATCTTCATCCTGCTGTTCCTCATCCCGTTCGTGATCACCGTGGCTTGT 923

Qy        301 TyrAlaAla------------IleIleArgThrLeuAsnAlaTyrAsp---HisArgTrp 315
              |||   |||            ::::::||||||      |||:::         :::
Db        924 TACACGGCCACCATCCTCAAGCTGTTGCGCACGGAGGAGGCGCACGGCCGGGAGCAGCGG 983

Qy        316 LeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 335
                  |||      :::   ::::::|||:::  |||      ||||||||||||
Db        984 AGCGCCGCGGTGGGCCTGGCCGCGGTGGTCTTGCTGGCCTTTGTCACCTGCTTCGCCCCC 1043

Qy        336 SerAsnIleIleLeuIleIleHis---HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
              :::|||   :::|||:::   |||      :::      :::|||      :::
Db       1044 AACAACTTCGTGCTCCTGGCGCACATCGTGAGCCGCCTGTTCTACGGCAAGAGC------ 1097

Qy        355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
                 |||   :::|||   ::::::|||||||||      ||||||:::||||||||||||
Db       1098 TACTACCACGTGTACAAGCTCAGCCTGTGTCTCAGCTGCCTCAACAACTGTCTGGACCCG 1157

Qy        375 PheLeuTyrPheLeuMetSerLys 382
              |||:::|||:::      |||:::
Db       1158 TTTGTTTATTACTTTGCGTCCCGG 1181
```

Search completed: June 29, 2003, 08:19:28
Job time : 247.736 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:29:54 ; Search time 2269.78 Seconds
(without alignments)
5231.316 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIKM.....AYLTKXXNDLREQGQPSQRT 408

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09208629/runat_24062003_101625_12922/app_query.fasta_1.1166
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629_@CGN_1_1_2827_@runat_24062003_101625_12922 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2142	99.3	1224	6	AR070426	AR070426 Sequence
2	2142	99.3	1830	9	HSU92971	U92971 Human prote
3	1960.5	90.9	9273	9	AF374726	AF374726 Homo sapi
c 4	1960.5	90.9	101887	9	AC026725	AC026725 Homo sapi
5	1960.5	90.9	131370	9	AC026706	AC026706 Homo sapi
6	1947	90.3	1102	6	AR070427	AR070427 Sequence
7	1443.5	66.9	1224	6	AR070424	AR070424 Sequence
8	1443.5	66.9	2409	10	MMU92972	U92972 Mus musculu
9	1424	66.0	1505	10	AF310076	AF310076 Rattus no
c 10	1367.5	63.4	235097	2	AC110877	AC110877 Mus muscu
11	959	44.5	1124	6	AR070425	AR070425 Sequence
12	606.5	28.1	2713	10	MMPAR2MR	Z48043 M.musculus
13	606.5	28.1	2732	6	AR012639	AR012639 Sequence
14	606.5	28.1	2732	6	AR171259	AR171259 Sequence
15	606.5	28.1	2732	6	I87850	I87850 Sequence 60
16	605.5	28.1	1475	6	AR012637	AR012637 Sequence
17	605.5	28.1	1475	6	AR171257	AR171257 Sequence
18	605.5	28.1	1475	6	I42454	I42454 Sequence 1
19	605.5	28.1	1475	6	I87848	I87848 Sequence 1
20	605.5	28.1	2772	10	BC025432	BC025432 Mus muscu
21	603.5	28.0	1477	10	MMPROACR2	Z35158 M.musculus
c 22	603.5	28.0	199227	2	AC110876	AC110876 Mus muscu
23	599.5	27.8	1680	5	XLU09632	U09632 Xenopus lae
24	595	27.6	1428	10	RNU61373	U61373 Rattus norv
25	592.5	27.5	2813	9	BC012453	BC012453 Homo sapi
c 26	592.5	27.5	88955	2	AC128225	AC128225 Rattus no
c 27	592.5	27.5	151166	2	AC130634	AC130634 Rattus no
28	591.5	27.4	2876	9	BC018130	BC018130 Homo sapi
29	586.5	27.2	1451	9	HSU34038	U34038 Human prote
30	585.5	27.1	1124	9	HSU36753	U36753 Human prote
31	585.5	27.1	1255	6	AR012638	AR012638 Sequence
32	585.5	27.1	1255	6	AR171258	AR171258 Sequence
33	585.5	27.1	1255	6	I42455	I42455 Sequence 3
34	585.5	27.1	1255	6	I87849	I87849 Sequence 3
35	585.5	27.1	1289	9	HSPAR2B	Z49994 H.sapiens p
36	585.5	27.1	18351	9	AF400075	AF400075 Homo sapi
37	585.5	27.1	52368	9	AC010621	AC010621 Homo sapi
c 38	585.5	27.1	184536	2	AC068682	AC068682 Homo sapi
39	585.5	27.1	204485	2	AC114962	AC114962 Homo sapi
40	572.5	26.5	1414	6	AR012640	AR012640 Sequence
41	572.5	26.5	1414	6	AR171260	AR171260 Sequence
42	572.5	26.5	1414	6	I87851	I87851 Sequence 62
c 43	555.5	25.8	153451	2	AC079576	AC079576 Mus muscu
c 44	554.5	25.7	118914	2	AC079554	AC079554 Mus muscu
45	543	25.2	1694	10	AF269246	AF269246 Rattus no

ALIGNMENTS

RESULT 1

```
AR070426
LOCUS       AR070426                1224 bp    DNA     linear   PAT 18-FEB-2000
DEFINITION  Sequence 4 from patent US 5892014.
ACCESSION   AR070426
VERSION     AR070426.1  GI:7221314
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1224)
  AUTHORS   Coughlin,S.R., Ishihara,H. and Connolly,A.
  TITLE     DNA encoding a protease-activated receptor 3
  JOURNAL   Patent: US 5892014-A 4 06-APR-1999;
FEATURES             Location/Qualifiers
     source          1. .1224
                     /organism="unknown"
BASE COUNT      303 a    316 c    221 g    384 t
ORIGIN

Alignment Scores:
Pred. No.:                2.01e-165      Length:          1224
Score:                    2142.00        Matches:         405
Percent Similarity:       99.26%         Conservative:    0
Best Local Similarity:    99.26%         Mismatches:      3
Query Match:              99.30%         Indels:          0
DB:                       6              Gaps:            0

US-09-208-629F-6 (1-408) x AR070426 (1-1224)

Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
              |||||||||||||||||||||||    |||||||||||||||||||||||||||||||||
Db          1 TGCTCCATGATTTTACAGATTTCATAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 60

Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 120

Qy         41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 180

Qy         61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 240

Qy         81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 300

Qy        101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 360

Qy        121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTCTTC 420

Qy        141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTTCTT 480

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTATTT 540

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 600

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTAC 660

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACAGTT 720

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 780

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTCTAT 840

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTACTGC 900

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 960

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATTCTT 1020

Qy        341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATATCTC 1080

Qy        361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTCATG 1140

Qy        381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
              |||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db       1141 TCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGAGAA 1200

Qy        401 GlnGlyGlnProSerGlnArgThr 408
              |||||||||||||||||||||||||
Db       1201 CAAGGACAGCCATCACAGAGAACG 1224

RESULT 2
HSU92971
LOCUS       HSU92971                1830 bp    mRNA    linear   PRI 16-APR-1997
DEFINITION  Human protease-activated receptor 3 (PAR3) mRNA, complete cds.
ACCESSION   U92971
VERSION     U92971.1  GI:1938374
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1830)
  AUTHORS   Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
            Timmons,C., Tram,T. and Coughlin,S.R.
  TITLE     Protease-activated receptor 3 is a second thrombin receptor in
            humans
  JOURNAL   Nature 386 (6624), 502-506 (1997)
  MEDLINE   97242411
   PUBMED   9087410
REFERENCE   2  (bases 1 to 1830)
  AUTHORS   Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
            Timmons,C., Tram,T. and Coughlin,S.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-1997) CVRI, UCSF, 3rd and Parnassus, San
            Francisco, CA 94143, USA
FEATURES             Location/Qualifiers
     source          1. .1830
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .1830
                     /gene="PAR3"
     CDS             145. .1269
                     /gene="PAR3"
```

```
                     /note="thrombin receptor; coagulation protease"
                     /codon_start=1
                     /product="protease-activated receptor 3"
                     /protein_id="AAC51218.1"
                     /db_xref="GI:1938375"
                     /translation="MKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAP
                     PNSFEEFPPSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYL
                     LVFVVGVPANAVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV
                     FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVW
                     ATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFV
                     LIIYCYAAIIRTLNAYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTD
                     GLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTK"
BASE COUNT      473 a    464 c    337 g    556 t
ORIGIN

Alignment Scores:
Pred. No.:              3.2e-165        Length:         1830
Score:                  2142.00         Matches:        405
Percent Similarity:     99.26%          Conservative:   0
Best Local Similarity:  99.26%          Mismatches:     3
Query Match:            99.30%          Indels:         0
DB:                     9               Gaps:           0

US-09-208-629F-6 (1-408) x HSU92971 (1-1830)

Qy          1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db         88 TGCTCCATGATTTTACAGATTTCATAACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 147

Qy         21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 AAAGCCCTCATCTTTGCAGCTGCTGGCCTCCTGCTTCTGTTGCCCACTTTTTGTCAGAGT 207

Qy         41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTTCGT 267

Qy         61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 GGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACAGGA 327

Qy         81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTGAAA 387

Qy        101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 AATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATCTAC 447

Qy        121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTCTTC 507

Qy        141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 AGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTTCTT 567

Qy        161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTATTT 627

Qy        181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 687

Qy        201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 CTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTAC 747

Qy        221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACAGTT 807

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 867

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 ATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTCTAT 927

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTACTGC 987

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 1047

Qy        321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 GCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATTCTT 1107

Qy        341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATATCTC 1167

Qy        361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 ATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTCATG 1227

Qy        381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
              ||||||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db       1228 TCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGAGAA 1287

Qy        401 GlnGlyGlnProSerGlnArgThr 408
              |||||||||||||||||||||||||
Db       1288 CAAGGACAGCCATCACAGAGAACG 1311

RESULT 3
AF374726
LOCUS       AF374726               9273 bp    DNA     linear   PRI 01-JUN-2001
DEFINITION  Homo sapiens coagulation factor II receptor-like 2 (F2RL2) gene,
            complete cds.
ACCESSION   AF374726
VERSION     AF374726.2  GI:14278710
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9273)
  AUTHORS   Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
            Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-2001) Molecular Biotechnology, University of
            Washington, 1705 NE Pacific, Seattle, WA 98195, USA
REFERENCE   2  (bases 1 to 9273)
  AUTHORS   Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
            Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2001) Molecular Biotechnology, University of
            Washington, 1705 NE Pacific, Seattle, WA 98195, USA
  REMARK    Sequence update by submitter
COMMENT     On Jun 1, 2001 this sequence version replaced gi:13991934.
            To cite this work please use:  SeattleSNPs. NHLBI Program for
            Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.mbt.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .9273
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     repeat_region   120. .221
                     /rpt_family="L1MB3"
                     /rpt_type=dispersed
     variation       419
```

```
                /frequency="0.29"
                /replace="G"
variation       483
                /frequency="0.02"
                /replace="T"
variation       506
                /frequency="0.32"
                /replace="G"
variation       507
                /frequency="0.25"
                /replace="C"
variation       512
                /frequency="0.24"
                /replace="A"
variation       535
                /frequency="0.24"
                /replace="T"
variation       937
                /frequency="0.02"
                /replace="C"
gene            969. .7322
                /gene="F2RL2"
mRNA            join(969. .1176,5701. .7322)
                /gene="F2RL2"
                /product="coagulation factor II receptor-like 2"
variation       1007
                /gene="F2RL2"
                /frequency="0.02"
                /replace="G"
CDS             join(1113. .1176,5701. .6761)
                /gene="F2RL2"
                /note="thrombin receptor-like 2"
                /codon_start=1
                /product="coagulation factor II receptor-like 2"
                /protein_id="AAK51564.1"
                /db_xref="GI:13991935"
                /translation="MKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAP
                PNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYL
                LVFVVGVPANAVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV
                FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVW
                ATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFV
                LIIYCYAAIIRTLNAYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTD
                GLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTK"
variation       1156
                /gene="F2RL2"
                /frequency="0.01"
                /replace="C"
variation       1430
                /gene="F2RL2"
                /frequency="0.06"
                /replace="G"
variation       2055
                /gene="F2RL2"
                /frequency="0.03"
                /replace="T"
variation       2120
                /gene="F2RL2"
                /frequency="0.01"
                /replace="T"
variation       2136
                /gene="F2RL2"
                /frequency="0.01"
                /replace="G"
variation       2168
                /gene="F2RL2"
                /frequency="0.02"
                /replace="T"
variation       2174
                /gene="F2RL2"
                /frequency="0.01"
                /replace="-"
repeat_region   2250. .2441
                /rpt_family="MER107"
                /rpt_type=dispersed
variation       2448
                /gene="F2RL2"
                /frequency="0.16"
                /replace="T"
variation       2535
                /gene="F2RL2"
                /frequency="0.38"
                /replace="T"
variation       2800
                /gene="F2RL2"
                /frequency="0.04"
                /replace="A"
variation       2866
                /gene="F2RL2"
                /frequency="0.35"
                /replace="T"
variation       3116
                /gene="F2RL2"
                /frequency="0.03"
                /replace="A"
variation       3135
                /gene="F2RL2"
                /frequency="0.02"
                /replace="A"
variation       3281
                /gene="F2RL2"
                /frequency="0.04"
                /replace="A"
variation       3386
                /gene="F2RL2"
                /frequency="0.02"
                /replace="A"
repeat_region   3543. .4180
                /rpt_family="L2"
                /rpt_type=dispersed
variation       3553
                /gene="F2RL2"
                /frequency="0.02"
                /replace="T"
variation       3565
                /gene="F2RL2"
                /frequency="0.08"
                /replace="T"
variation       3610
                /gene="F2RL2"
                /frequency="0.01"
                /replace="-"
variation       3694
                /gene="F2RL2"
                /frequency="0.07"
                /replace="C"
variation       3734
                /gene="F2RL2"
                /frequency="0.32"
                /replace="T"
variation       3850
                /gene="F2RL2"
                /frequency="0.02"
                /replace="A"
variation       3976
                /gene="F2RL2"
                /frequency="0.60"
                /replace="-"
repeat_region   4181. .4526
                /rpt_family="MLT1D"
                /rpt_type=dispersed
variation       4223. .4226
                /gene="F2RL2"
                /frequency="0.02"
                /replace="-"
variation       4264
                /gene="F2RL2"
```

```
                     /frequency="0.32"
                     /replace="C"
     variation       4305
                     /gene="F2RL2"
                     /frequency="0.35"
                     /replace="G"
     variation       4464
                     /gene="F2RL2"
                     /frequency="0.01"
                     /replace="G"
     repeat_region   4527..4827
                     /rpt_family="AluY"
                     /rpt_type=dispersed
     variation       4556
                     /gene="F2RL2"
                     /frequency="0.03"
                     /replace="A"
     variation       4580
                     /gene="F2RL2"

Alignment Scores:
Pred. No.:              1.2e-149        Length:         9273
Score:                  1960.50         Matches:        375
Percent Similarity:     95.49%          Conservative:   6
Best Local Similarity:  93.98%          Mismatches:     13
Query Match:            90.89%          Indels:         5
DB:                     9               Gaps:           1

US-09-208-629F-6 (1-408) x AF374726 (1-9273)

Qy         12 ArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeu 31
              |||:::      |||:::||||||             |||   :::      |||
Db       5620 AGAAATTATTTACAAATAATTAAACAC---------ACTTTCTTTTCTCTTGAACTATAC 5670

Qy         32 LeuLeuLeuProThrPheCysGln--SerGlyMetGluAsnAspThrAsnAsnLeuAlaL 51
              :::   :::   ||||||   |||  :::||||||||||||||||||||||||||||||
Db       5671 ATTTATGTTAATACTTTCTTTCAATTACAGGCATGGAAAATGATACAAACAACTTGGCAA 5730

Qy         51 ysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheP 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5731 AGCCAACCTTACCCATTAAGACCTTTCGTGGAGCTCCCCCAAATTCTTTTGAAGAGTTCC 5790

Qy         71 roPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProG 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5791 CCTTTTCTGCCTTGGAAGGCTGGACAGGAGCCACGATTACTGTAAAAATTAAGTGCCCTG 5850

Qy         91 luGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerL 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5851 AAGAAAGTGCTTCACATCTCCATGTGAAAAATGCTACCATGGGGTACCTGACCAGCTCCT 5910

Qy        111 euSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaA 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5911 TAAGTACTAAACTGATACCTGCCATCTACCTCCTGGTGTTTGTAGTTGGTGTCCCGGCCA 5970

Qy        131 snAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheT 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5971 ATGCTGTGACCCTGTGGATGCTTTTCTTCAGGACCAGATCCATCTGTACCACTGTATTCT 6030

Qy        151 yrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaT 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6031 ACACCAACCTGGCCATTGCAGATTTTCTTTTTTGTGTTACATTGCCCTTTAAGATAGCTT 6090

Qy        171 yrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValI 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6091 ATCATCTCAATGGGAACAACTGGGTATTTGGAGAGGTCCTGTGCCGGGCCACCACAGTCA 6150

Qy        191 lePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgT 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6151 TCTTCTATGGCAACATGTACTGCTCCATTCTGCTCCTTGCCTGCATCAGCATCAACCGCT 6210

Qy        211 yrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuV 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6211 ACCTGGCCATCGTCCATCCTTTCACCTACCGGGGCCTGCCCAAGCACACCTATGCCTTGG 6270

Qy        231 alThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuL 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6271 TAACATGTGGACTGGTGTGGGCAACAGTTTTCTTATATATGCTGCCATTTTTCATACTGA 6330

Qy        251 ysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrC 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6331 AGCAGGAATATTATCTTGTTCAGCCAGACATCACCACCTGCCATGATGTTCACAACACTT 6390

Qy        271 ysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheL 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6391 GCGAGTCCTCATCTCCCTTCCAACTCTATTACTTCATCTCCTTGGCATTCTTTGGATTCT 6450

Qy        291 euIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaT 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6451 TAATTCCATTTGTGCTTATCATCTACTGCTATGCAGCCATCATCCGGACACTTAATGCAT 6510

Qy        311 yrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrI 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6511 ACGATCATAGATGGTTGTGGTATGTTAAGGCGAGTCTCCTCATCCTTGTGATTTTTACCA 6570

Qy        331 leCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnA 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6571 TTTGCTTTGCTCCAAGCAATATTATTCTTATTATTCACCATGCTAACTACTACTACAACA 6630

Qy        351 snThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerC 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6631 ACACTGATGGCTTATATTTTATATATCTCATAGCTTTGTGCCTGGGTAGTCTTAATAGTT 6690

Qy        371 ysLeuAspProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrL 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6691 GCTTAGATCCATTCCTTTATTTTCTCATGTCAAAAACCAGAAATCACTCCACTGCTTACC 6750

Qy        391 euThrLys******AsnAspLeuArgGluGlnGlyGlnProSerGlnArgThr 408
              ||||||||      ||||||||||||||||||||||||||||||||||||||
Db       6751 TTACAAAATAGTGAAATGATCTTAGAGAACAAGGACAGCCATCACAGAGAACG 6803

RESULT 4
AC026725/c
LOCUS       AC026725              101887 bp    DNA     linear   PRI 21-JUL-2001
DEFINITION  Homo sapiens chromosome 5 clone CTD-2236F14, complete sequence.
ACCESSION   AC026725
VERSION     AC026725.7  GI:14993696
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 101887)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 101887)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 101887)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 101887)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 21, 2001 this sequence version replaced gi:14572130.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
```

```
           Estimated Total Number of Errors is 0.1.
           STS Content:
           SHGC-102125 G56565
           SHGC-101620 G56147
           SHGC-83998 G53567.
FEATURES             Location/Qualifiers
     source          1. .101887
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2236F14"
BASE COUNT    29200 a  20058 c  21105 g  31524 t
ORIGIN

Alignment Scores:
Pred. No.:              1.88e-148      Length:         101887
Score:                  1960.50        Matches:        375
Percent Similarity:     95.49%         Conservative:   6
Best Local Similarity:  93.98%         Mismatches:     13
Query Match:            90.89%         Indels:         5
DB:                     9              Gaps:           1

US-09-208-629F-6 (1-408) x AC026725 (1-101887)

Qy         12 ArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeu 31
              |||:::     |||:::||||||               |||   :::     |||
Db      93557 AGAAATTATTTACAAATAATTAAACAC---------ACTTTCTTTTCTCTTGAACTATAC 93507

Qy         32 LeuLeuLeuProThrPheCysGln--SerGlyMetGluAsnAspThrAsnAsnLeuAlaL 51
              :::   :::   ||||||  |||  :::||||||||||||||||||||||||||||||||
Db      93506 ATTTATGTTAATACTTTCTTTCAATTACAGGCATGGAAAATGATACAAACAACTTGGCAA 93447

Qy         51 ysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheP 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93446 AGCCAACCTTACCCATTAAGACCTTTCGTGGAGCTCCCCCAAATTCTTTTGAAGAGTTCC 93387

Qy         71 roPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProG 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93386 CCTTTTCTGCCTTGGAAGGCTGGACAGGAGCCACGATTACTGTAAAAATTAAGTGCCCTG 93327

Qy         91 luGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerL 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93326 AAGAAAGTGCTTCACATCTCCATGTGAAAAATGCTACCATGGGGTACCTGACCAGCTCCT 93267

Qy        111 euSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaA 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93266 TAAGTACTAAACTGATACCTGCCATCTACCTCCTGGTGTTTGTAGTTGGTGTCCCGGCCA 93207

Qy        131 snAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheT 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93206 ATGCTGTGACCCTGTGGATGCTTTTCTTCAGGACCAGATCCATCTGTACCACTGTATTCT 93147

Qy        151 yrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaT 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93146 ACACCAACCTGGCCATTGCAGATTTTCTTTTTTGTGTTACATTGCCCTTTAAGATAGCTT 93087

Qy        171 yrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValI 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93086 ATCATCTCAATGGGAACAACTGGGTATTTGGAGAGGTCCTGTGCCGGGCCACCACAGTCA 93027

Qy        191 lePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgT 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93026 TCTTCTATGGCAACATGTACTGCTCCATTCTGCTCCTTGCCTGCATCAGCATCAACCGCT 92967

Qy        211 yrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuV 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92966 ACCTGGCCATCGTCCATCCTTTCACCTACCGGGGCCTGCCCAAGCACACCTATGCCTTGG 92907

Qy        231 alThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuL 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92906 TAACATGTGGACTGGTGTGGGCAACAGTTTTCTTATATATGCTGCCATTTTTCATACTGA 92847

Qy        251 ysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrC 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92846 AGCAGGAATATTATCTTGTTCAGCCAGACATCACCACCTGCCATGATGTTCACAACACTT 92787

Qy        271 ysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheL 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92786 GCGAGTCCTCATCTCCCTTCCAACTCTATTACTTCATCTCCTTGGCATTCTTTGGATTCT 92727

Qy        291 euIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaT 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92726 TAATTCCATTTGTGCTTATCATCTACTGCTATGCAGCCATCATCCGGACACTTAATGCAT 92667

Qy        311 yrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrI 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92666 ACGATCATAGATGGTTGTGGTATGTTAAGGCGAGTCTCCTCATCCTTGTGATTTTTACCA 92607

Qy        331 leCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnA 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92606 TTTGCTTTGCTCCAAGCAATATTATTCTTATTATTCACCATGCTAACTACTACTACAACA 92547

Qy        351 snThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerC 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92546 ACACTGATGGCTTATATTTTATATATCTCATAGCTTTGTGCCTGGGTAGTCTTAATAGTT 92487

Qy        371 ysLeuAspProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrL 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92486 GCTTAGATCCATTCCTTTATTTTCTCATGTCAAAAACCAGAAATCACTCCACTGCTTACC 92427

Qy        391 euThrLys******AsnAspLeuArgGluGlnGlyGlnProSerGlnArgThr 408
              ||||||||      ||||||||||||||||||||||||||||||||||||||
Db      92426 TTACAAAATAGTGAAATGATCTTAGAGAACAAGGACAGCCATCACAGAGAACG 92374

RESULT 5
AC026706
LOCUS       AC026706               131370 bp    DNA     linear   PRI 21-JUN-2001
DEFINITION  Homo sapiens chromosome 5 clone CTD-2068N7, complete sequence.
ACCESSION   AC026706
VERSION     AC026706.5  GI:14518404
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 131370)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 131370)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 131370)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jun 21, 2001 this sequence version replaced gi:13677033.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.6% of Sequence;
            Estimated Total Number of Errors is 0.3.
            STS Content:
            WI-11362 G24506.
FEATURES             Location/Qualifiers
     source          1. .131370
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2068N7"
BASE COUNT    40462 a  26452 c  25592 g  38864 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              2.52e-148       Length:         131370
Score:                  1960.50         Matches:        375
Percent Similarity:     95.49%          Conservative:   6
Best Local Similarity:  93.98%          Mismatches:     13
Query Match:            90.89%          Indels:         5
DB:                     9               Gaps:           1

US-09-208-629F-6 (1-408) x AC026706 (1-131370)

Qy         12 ArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeu 31
              |||:::     |||:::||||||               |||   :::      |||
Db      60121 AGAAATTATTTACAAATAATTAAACAC---------ACTTTCTTTTCTCTTGAACTATAC 60171

Qy         32 LeuLeuLeuProThrPheCysGln--SerGlyMetGluAsnAspThrAsnAsnLeuAlaL 51
              :::   :::   ||||||   |||  :::||||||||||||||||||||||||||||||
Db      60172 ATTTATGTTAATACTTTCTTTCAATTACAGGCATGGAAAATGATACAAACAACTTGGCAA 60231

Qy         51 ysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheP 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60232 AGCCAACCTTACCCATTAAGACCTTTCGTGGAGCTCCCCCAAATTCTTTTGAAGAGTTCC 60291

Qy         71 roPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProG 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60292 CCTTTTCTGCCTTGGAAGGCTGGACAGGAGCCACGATTACTGTAAAAATTAAGTGCCCTG 60351

Qy         91 luGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerL 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60352 AAGAAAGTGCTTCACATCTCCATGTGAAAAATGCTACCATGGGGTACCTGACCAGCTCCT 60411

Qy        111 euSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaA 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60412 TAAGTACTAAACTGATACCTGCCATCTACCTCCTGGTGTTTGTAGTTGGTGTCCCGGCCA 60471

Qy        131 snAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheT 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60472 ATGCTGTGACCCTGTGGATGCTTTTCTTCAGGACCAGATCCATCTGTACCACTGTATTCT 60531

Qy        151 yrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaT 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60532 ACACCAACCTGGCCATTGCAGATTTTCTTTTTTGTGTTACATTGCCCTTTAAGATAGCTT 60591

Qy        171 yrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValI 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60592 ATCATCTCAATGGGAACAACTGGGTATTTGGAGAGGTCCTGTGCCGGGCCACCACAGTCA 60651

Qy        191 lePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgT 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60652 TCTTCTATGGCAACATGTACTGCTCCATTCTGCTCCTTGCCTGCATCAGCATCAACCGCT 60711

Qy        211 yrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuV 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60712 ACCTGGCCATCGTCCATCCTTTCACCTACCGGGGCCTGCCCAAGCACACCTATGCCTTGG 60771

Qy        231 alThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuL 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60772 TAACATGTGGACTGGTGTGGGCAACAGTTTTCTTATATATGCTGCCATTTTTCATACTGA 60831

Qy        251 ysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrC 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60832 AGCAGGAATATTATCTTGTTCAGCCAGACATCACCACCTGCCATGATGTTCACAACACTT 60891

Qy        271 ysGluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheL 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60892 GCGAGTCCTCATCTCCCTTCCAACTCTATTACTTCATCTCCTTGGCATTCTTTGGATTCT 60951

Qy        291 euIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaT 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60952 TAATTCCATTTGTGCTTATCATCTACTGCTATGCAGCCATCATCCGGACACTTAATGCAT 61011

Qy        311 yrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrI 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61012 ACGATCATAGATGGTTGTGGTATGTTAAGGCGAGTCTCCTCATCCTTGTGATTTTTACCA 61071

Qy        331 leCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnA 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61072 TTTGCTTTGCTCCAAGCAATATTATTCTTATTATTCACCATGCTAACTACTACTACAACA 61131

Qy        351 snThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerC 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61132 ACACTGATGGCTTATATTTTATATATCTCATAGCTTTGTGCCTGGGTAGTCTTAATAGTT 61191

Qy        371 ysLeuAspProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrL 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61192 GCTTAGATCCATTCCTTTATTTTCTCATGTCAAAAACCAGAAATCACTCCACTGCTTACC 61251

Qy        391 euThrLys******AsnAspLeuArgGluGlnGlyGlnProSerGlnArgThr 408
              |||||||      ||||||||||||||||||||||||||||||||||||
Db      61252 TTACAAAATAGTGAAATGATCTTAGAGAACAAGGACAGCCATCACAGAGAACG 61304

RESULT 6
AR070427
LOCUS       AR070427                1102 bp    DNA     linear   PAT 18-FEB-2000
DEFINITION  Sequence 5 from patent US 5892014.
ACCESSION   AR070427
VERSION     AR070427.1  GI:7221315
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1102)
  AUTHORS   Coughlin,S.R., Ishihara,H. and Connolly,A.
  TITLE     DNA encoding a protease-activated receptor 3
  JOURNAL   Patent: US 5892014-A 5 06-APR-1999;
FEATURES             Location/Qualifiers
     source          1. .1102
                     /organism="unknown"
BASE COUNT      276 a    286 c    195 g    345 t
ORIGIN

Alignment Scores:
Pred. No.:              1.3e-149        Length:         1102
Score:                  1947.00         Matches:        364
Percent Similarity:     99.46%          Conservative:   1
Best Local Similarity:  99.18%          Mismatches:     2
Query Match:            90.26%          Indels:         0
DB:                     6               Gaps:           0

US-09-208-629F-6 (1-408) x AR070427 (1-1102)

Qy         40 SerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACAGGCATGGAAAATGATACAAACAACTTGGCAAAGCCAACCTTACCCATTAAGACCTTT 60

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThr 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGGAGCTCCCCCAAATTCTTTTGAAGAGTTCCCCTTTTCTGCCTTGGAAGGCTGGACA 120

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTGCTTCACATCTCCATGTG 180

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAAATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACTAAACTGATACCTGCCATC 240

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACCTCCTGGTGTTTGTAGTTGGTGTCCCGGCCAATGCTGTGACCCTGTGGATGCTTTTC 300

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCAGGACCAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGCAGATTTT 360

Qy        160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
```

```
Db        361 CTTTTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACTGGGTA 420

Qy        180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199

Db        421 TTTGGAGAGGTCCTGTGCCGGGCCACCACAGTCATCTTCTATGGCAACATGTACTGCTCC 480

Qy        200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219

Db        481 ATTCTGCTCCTTGCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACC 540

Qy        220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239

Db        541 TACCGGGGCCTGCCCAAGCACACCTATGCCTTGGTAACATGTGGACTGGTGTGGGCAACA 600

Qy        240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259

Db        601 GTTTTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCA 660

Qy        260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279

Db        661 GACATCACCACCTGCCATGATGTTCACAACACTTGCGAGTCCTCATCTCCCTTCCAACTC 720

Qy        280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299

Db        721 TATTACTTCATCTCCTTGGCATTCTTTGGATTCTTAATTCCATTTGTGCTTATCATCTAC 780

Qy        300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrVal 319

Db        781 TGCTATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTT 840

Qy        320 LysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIle 339

Db        841 AAGGCGAGTCTCCTCATCCTTGTGATTTTTACCATTTGCTTTGCTCCAAGCAATATTATT 900

Qy        340 LeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyr 359

Db        901 CTTATTATTCACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTTTATATAT 960

Qy        360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379

Db        961 CTCATAGCTTTGTGCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTTCTC 1020

Qy        380 MetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArg 399

Db       1021 ATGTCAAAAACCAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAAATGATCTTAGA 1080

Qy        400 GluGlnGlyGlnProSerGln 406

Db       1081 GAACAAGGACAGCCATCACAG 1101

RESULT 7
AR070424
LOCUS       AR070424                1224 bp    DNA     linear   PAT 18-FEB-2000
DEFINITION  Sequence 1 from patent US 5892014.
ACCESSION   AR070424
VERSION     AR070424.1  GI:7221312
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1224)
  AUTHORS   Coughlin,S.R., Ishihara,H. and Connolly,A.
  TITLE     DNA encoding a protease-activated receptor 3
  JOURNAL   Patent: US 5892014-A 1 06-APR-1999;
FEATURES             Location/Qualifiers
     source          1. .1224
                     /organism="unknown"
BASE COUNT      299 a    341 c    239 g    345 t
ORIGIN

Alignment Scores:
Pred. No.:              1.33e-108      Length:        1224

Score:                  1443.50        Matches:       264
Percent Similarity:     83.74%         Conservative:  45
Best Local Similarity:  71.54%         Mismatches:    59
Query Match:            66.92%         Indels:        1
DB:                     6              Gaps:          1

US-09-208-629F-6 (1-408) x AR070424 (1-1224)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33

Db         33 GGGTCTCAGGACATCAAGATGAAAATCCTTATCTTGGTTGCAGCTGGGCTGCTGTTTCTG 92

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53

Db         93 CCAGTCACTGTTTGCCAAAGTGGCATA---AATGTTTCAGACAACTCAGCAAAGCCAACC 149

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSer 73

Db        150 TTAACTATTAAGAGTTTTAATGGGGGTCCCCAAAATACCTTTGAAGAATTCCCACTTTCT 209

Qy         74 AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93

Db        210 GACATAGAGGGCTGGACAGGAGCCACCACAACTATAAAAGCGGAGTGTCCCGAGGACAGT 269

Qy         94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThr 113

Db        270 ATTTCAACTCTCCACGTGAATAATGCTACCATAGGATACCTGAGAAGTTCCTTAAGTACC 329

Qy        114 LysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaVal 133

Db        330 CAAGTGATACCTGCCATCTATATCCTGCTGTTTGTGGTTGGTGTACCATCCAACATCGTG 389

Qy        134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153

Db        390 ACCCTGTGGAAACTCTCCTTAAGGACCAAATCCATCAGTCTGGTCATCTTTCACACCAAC 449

Qy        154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173

Db        450 CTGGCCATCGCAGATCTCCTTTTCTGTGTCACACTGCCATTTAAGATCGCCTACCATCTC 509

Qy        174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193

Db        510 AATGGCAACAACTGGGTATTTGGCGAGGTCATGTGCCGGATCACCACGGTCGTTTTCTAC 569

Qy        194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213

Db        570 GGCAACATGTACTGCGCTATCCTGATCCTCACTTGCATGGGCATCAACCGCTACCTGGCC 629

Qy        214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233

Db        630 ACGGCTCACCCTTTCACATACCAGAAGCTGCCCAAACGCAGCTTCTCCTTGCTCATGTGT 689

Qy        234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253

Db        690 GGCATAGTGTGGGTCATGGTTTTCTTATACATGCTGCCCTTTGTCATCCTGAAGCAGGAG 749

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273

Db        750 TACCACCTCGTCCACTCAGAGATCACCACCTGCCACGATGTCGTCGACGCGTGCGAGTCC 809

Qy        274 SerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIlePro 293

Db        810 CCATCATCCTTCCGATTCTACTACTTCGTCTCCTTAGCATTCTTTGGGTTCCTCATCCCG 869

Qy        294 PheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHis 313

Db        870 TTTGTGATCATCATCTTCTGTTACACGACTCTCATCCACAAACTTAAATCAAAGGATCGG 929

Qy        314 ArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPhe 333

Db        930 ATATGGCTGGGCTACATCAAGGCCGTCCTCCTCATCCTTGTGATTTTCACAATTTGCTTT 989

Qy        334 AlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAsp 353
```

```
Db        990 GCCCCCACCAACATCATACTCGTAATCCACCATGCCAACTACTACTACCACAATACCGAC 1049

Qy        354 GlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373

Db       1050 AGCTTGTACTTTATGTATCTTATTGCTCTGTGCCTGGGGAGCCTGAATAGCTGCCTAGAT 1109

Qy        374 ProPheLeuTyrPheLeuMetSerLys 382

Db       1110 CCATTCCTTTACTTTGTCATGTCGAAA 1136
```

```
RESULT 8
MMU92972
LOCUS       MMU92972                2409 bp    mRNA    linear   ROD 16-APR-1997
DEFINITION  Mus musculus protease-activated receptor 3 (PAR3) mRNA, complete
            cds.
ACCESSION   U92972
VERSION     U92972.1  GI:1938376
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2409)
  AUTHORS   Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
            Timmons,C., Tram,T. and Coughlin,S.R.
  TITLE     Protease-activated receptor 3 is a second thrombin receptor in
            humans
  JOURNAL   Nature 386 (6624), 502-506 (1997)
  MEDLINE   97242411
   PUBMED   9087410
REFERENCE   2  (bases 1 to 2409)
  AUTHORS   Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
            Timmons,C., Tram,T. and Coughlin,S.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-1997) CVRI, UCSF, 3rd and Parnassus, San
            Francisco, CA 94143, USA
FEATURES             Location/Qualifiers
     source          1. .2409
                     /organism="Mus musculus"
                     /strain="C57/B6"
                     /db_xref="taxon:10090"
     gene            1. .2409
                     /gene="PAR3"
     CDS             153. .1262
                     /gene="PAR3"
                     /note="thrombin receptor; coagulation protease"
                     /codon_start=1
                     /product="protease-activated receptor 3"
                     /protein_id="AAC53137.1"
                     /db_xref="GI:1938377"
                     /translation="MKILILVAAGLLFLPVTVCQSGINVSDNSAKPTLTIKSFNGGPQ
                     NTFEEFPLSDIEGWTGATTTIKAECPEDSISTLHVNNATIGYLRSSLSTQVIPAIYIL
                     LFVVGVPSNIVTLWKLSLRTKSISLVIFHTNLAIADLLFCVTLPFKIAYHLNGNNWVF
                     GEVMCRITTVVFYGNMYCAILILTCMGINRYLATAHPFTYQKLPKRSFSLLMCGIVWV
                     MVFLYMLPFVILKQEYHLVHSEITTCHDVVDACESPSSFRFYYFVSLAFFGFLIPFVI
                     IIFCYTTLIHKLKSKDRIWLGYIKAVLLILVIFTICFAPTNIILVIHHANYYYHNTDS
                     LYFMYLIALCLGSLNSCLDPFLYFVMSKVVDQLNP"
BASE COUNT      598 a    652 c    454 g    703 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              2.88e-108      Length:         2409
Score:                  1443.50        Matches:        264
Percent Similarity:     83.74%         Conservative:   45
Best Local Similarity:  71.54%         Mismatches:     59
Query Match:            66.92%         Indels:         1
DB:                     10             Gaps:           1

US-09-208-629F-6 (1-408) x MMU92972 (1-2409)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
```

```
Db        135 GGGTCTCAGGACATCAAGATGAAAATCCTTATCTTGGTTGCAGCTGGGCTGCTGTTTCTG 194

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53

Db        195 CCAGTCACTGTTTGCCAAAGTGGCATA---AATGTTTCAGACAACTCAGCAAAGCCAACC 251

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSer 73

Db        252 TTAACTATTAAGAGTTTTAATGGGGGTCCCCAAAATACCTTTGAAGAATTCCCACTTTCT 311

Qy         74 AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93

Db        312 GACATAGAGGGCTGGACAGGAGCCACCACAACTATAAAAGCGGAGTGTCCCGAGGACAGT 371

Qy         94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThr 113

Db        372 ATTTCAACTCTCCACGTGAATAATGCTACCATAGGATACCTGAGAAGTTCCTTAAGTACC 431

Qy        114 LysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaVal 133

Db        432 CAAGTGATACCTGCCATCTATATCCTGCTGTTTGTGGTTGGTGTACCATCCAACATCGTG 491

Qy        134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153

Db        492 ACCCTGTGGAAACTCTCCTTAAGGACCAAATCCATCAGTCTGGTCATCTTTCACACCAAC 551

Qy        154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173

Db        552 CTGGCCATCGCAGATCTCCTTTTCTGTGTCACACTGCCATTTAAGATCGCCTACCATCTC 611

Qy        174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193

Db        612 AATGGCAACAACTGGGTATTTGGCGAGGTCATGTGCCGGATCACCACGGTCGTTTTCTAC 671

Qy        194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213

Db        672 GGCAACATGTACTGCGCTATCCTGATCCTCACTTGCATGGGCATCAACCGCTACCTGGCC 731

Qy        214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233

Db        732 ACGGCTCACCCTTTCACATACCAGAAGCTGCCCAAACGCAGCTTCTCCTTGCTCATGTGT 791

Qy        234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253

Db        792 GGCATAGTGTGGGTCATGGTTTTCTTATACATGCTGCCCTTTGTCATCCTGAAGCAGGAG 851

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273

Db        852 TACCACCTCGTCCACTCAGAGATCACCACCTGCCACGATGTCGTCGACGCGTGCGAGTCC 911

Qy        274 SerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIlePro 293

Db        912 CCATCATCCTTCCGATTCTACTACTTCGTCTCCTTAGCATTCTTTGGGTTCCTCATCCCG 971

Qy        294 PheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHis 313

Db        972 TTTGTGATCATCATCTTCTGTTACACGACTCTCATCCACAAACTTAAATCAAAGGATCGG 1031

Qy        314 ArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPhe 333

Db       1032 ATATGGCTGGGCTACATCAAGGCCGTCCTCCTCATCCTTGTGATTTTCACAATTTGCTTT 1091

Qy        334 AlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAsp 353

Db       1092 GCCCCCACCAACATCATACTCGTAATCCACCATGCCAACTACTACTACCACAATACCGAC 1151

Qy        354 GlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373

Db       1152 AGCTTGTACTTTATGTATCTTATTGCTCTGTGCCTGGGGAGCCTGAATAGCTGCCTAGAT 1211

Qy        374 ProPheLeuTyrPheLeuMetSerLys 382

Db       1212 CCATTCCTTTACTTTGTCATGTCGAAA 1238
```

```
RESULT 9
AF310076
LOCUS       AF310076               1505 bp    mRNA    linear   ROD 02-NOV-2001
DEFINITION  Rattus norvegicus protease activated receptor 3 mRNA, complete cds.
ACCESSION   AF310076
VERSION     AF310076.1  GI:16588398
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1505)
  AUTHORS   Chien,E.K., Marietti,S., Mendoza,J. and Phillippe,M.
  TITLE     Cloning of the rat protease activated receptor isoforms 3 and 4
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1505)
  AUTHORS   Chien,E.K., Marietti,S., Mendoza,J. and Phillippe,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2000) Obstetrics and Gynecology, University of
            Chicago, 5841 South Maryland Ave, MC2050, Chicago, IL 60637, USA
FEATURES             Location/Qualifiers
     source          1. .1505
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="spleen"
     CDS             63. .1169
                     /note="PAR3; G-protein coupled receptor"
                     /codon_start=1
                     /product="protease activated receptor 3"
                     /protein_id="AAL26789.1"
                     /db_xref="GI:16588399"
                     /translation="MEMKVLILVGVRLLFLPTTVCQSGMKHVSDNSALTAESFNGNEH
                     SFEEFPLSDIEGWTGATTTIKAKCPEESITTLHVNNATMGYLRSSLSTKVIPAIYILV
                     FVIGVPANIVTLWKLSSRTKSICLVIFHTNLAIADLLFCVTLPFKIAYHLNGNDWVFG
                     EVMCRVTTVAFYGNMYCAILILTCMGINRYLATVHPFTYRKLPKRNFTLLMCGVVWVM
                     VVLYMLPLAILKQEYHLVQPGITTCHDVHDTCESPLPFQFYYFVSLAFFGFLIPFVVS
                     VFCYTTLIHKLNAQDRKWLRYIKAVLLILVIFTICFAPTNIILIIHHANYYYSNTDSL
                     YFMYLIALCLGSLNSCLDPFLYFIMSKIVDQLTS"
BASE COUNT      359 a    394 c    316 g    436 t
ORIGIN

Alignment Scores:
Pred. No.:              6.48e-107      Length:         1505
Score:                  1424.00        Matches:        266
Percent Similarity:     80.05%         Conservative:   43
Best Local Similarity:  68.91%         Mismatches:     73
Query Match:            66.02%         Indels:         4
DB:                     10             Gaps:           2

US-09-208-629F-6 (1-408) x AF310076 (1-1505)

Qy         14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeu 33
              |||:::|||   ::::::||||||   ||||||                  ||||||   |||
Db         51 GGGTCTCAAGACATGGAGATGAAAGTCCTTATCTTGGTTGGGGGTTAGGCTGCTGTTTCTG 110

Qy         34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
                 |||   |||||||||||||||::::::   ::::::|||   |||
Db        111 CCGACCACTGTTTGCCAAAGTGGCATGAAACATGTTTCAGACAACTCAGCC--------- 161

Qy         54 LeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluProPheSer 73
              |||      ::::::|||   |||         :::|||||||||||||||||   |||
Db        162 TTAACTGCTGAGAGTTTTAATGGC---AACGAACATTCCTTTGAAGAATTCCCACTTTCT 218

Qy         74 AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93
                ::::||||||||||||||||||||   |||:::|||   ||||||||||||||||||
Db        219 GACATAGAGGGCTGGACAGGAGCCACCACAACTATAAAAGCGAAGTGTCCCGAAGAAAGC 278

Qy         94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThr 113
                ::: ||||||||| ||||||||||||||||||||   ||||||||||||||||
Db        279 ATTACAACTCTCCATGTGAATAATGCTACCATGGGATACCTGAGAAGTTCCTTAAGTACC 338

Qy        114 LysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaVal 133
              |||:::|||||||||||||||:::|||||||||||:::|||||||||||||||   |||
Db        339 AAAGTGATACCTGCCATCTACATCCTGGTGTTTGTGATTGGTGTACCAGCGAACATCGTG 398

Qy        134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153
              |||||||||   |||       ||||||:::|||||||||       :::|||:::||||||
Db        399 ACCCTGTGGAAACTCTCCTCAAGGACCAAATCCATCTGTCTGGTCATCTTTCACACCAAC 458

Qy        154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        459 CTGGCCATCGCGGATCTCCTTTTCTGTGTCACGCTGCCGTTTAAGATCGCCTACCATCTC 518

Qy        174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
              |||||||||:::|||||||||||||||||:::||||||   |||||||||   ||||||
Db        519 AATGGGAACGACTGGGTCTTTGGCGAGGTCATGTGCCGGGTCACCACGGTCGCTTTCTAC 578

Qy        194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
              ||||||||||||||||:::||||||:::|||   |||:::   ||||||||||||||||||
Db        579 GGCAACATGTATTGTGCTATTCTCATCCTCACCTGTATGGGCATCAACCGCTACCTGGCC 638

Qy        214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233
              ||||||||||||||||||||   |||||||||      :::   |||:::   |||
Db        639 ACGGTCCATCCCTTCACATACCGCAAGCTGCCCAAACGCAACTTCACTTTGCTCATGTGT 698

Qy        234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
              |||:::||||||       |||   ||||||||||||||      ||||||||||||||
Db        699 GGCGTGGTGTGGGTCATGGTGGTCTTATATATGTTGCCGTTGGCCATCCTCAAGCAGGAG 758

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
              |||:::|||||||||||||   ||||||||||||||||||||||:::|||||||||||||
Db        759 TACCATCTTGTCCAACCAGGGATCACCACCTGCCACGACGTCCACGACACATGCGAGTCC 818

Qy        274 SerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIlePro 293
                    ||||||||||   ||||||||||:::||||||||||||||||||||||||||||||
Db        819 CCGTTGCCCTTCCAGTTCTACTACTTCGTCTCCTTGGCCTTCTTTGGGTTCCTTATCCCA 878

Qy        294 PheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHis 313
              ||||||:::   ::::::||||||       :::|||      ||||||||   |||
Db        879 TTTGTGGTCAGTGTCTTCTGTTACACGACTCTCATCCACAAGCTTAACGCGCAGGATCGC 938

Qy        314 ArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPhe 333
              :::||||||   |||:::||||||   ||||||||||||||||||||||||||||||||||
Db        939 AAATGGCTGAGGTACATCAAGGCGGTTCTCCTCATCCTTGTGATTTTCACCATCTGCTTT 998

Qy        334 AlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAsp 353
              ||||||:::||||||||||||||||||||||||||||||||||||||||||:::||||||||
Db        999 GCCCCCACCAACATCATACTCATAATTCACCACGCCAACTACTACTACAGCAACACTGAT 1058

Qy        354 GlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
                 ||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AGCTTGTACTTTATGTATCTCATAGCTCTCTGCCTGGGCAGCCTGAATAGTTGCCTAGAC 1118

Qy        374 ProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys 393
              ||||||||||||||:::||||||||||      :::      |||:::      |||
Db       1119 CCATTCCTTTATTTTATCATGTCAAAAATTGTAGATCAGCTGACCTCGTAGTCAACAATG 1178

Qy        394 ******AsnAspLeuArg 399
                    :::      |||
Db       1179 GCGAGATCATCTTGGAGA 1196

RESULT 10
AC110877/c
LOCUS       AC110877               235097 bp    DNA     linear   HTG 16-FEB-2002
DEFINITION  Mus musculus clone RP23-345I21, WORKING DRAFT SEQUENCE, 3 unordered
            pieces.
ACCESSION   AC110877
VERSION     AC110877.1  GI:18693419
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE   1  (bases 1 to 235097)
  AUTHORS   Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and Cheng,J.F.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 235097)
  AUTHORS   Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and
            Cheng,J.-F.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2002) Berkeley PGA, Lawrence Berkeley National
            Laboratory, 1 Cyclotron rd., Berkeley, CA 94720, United States of
            America
COMMENT     Draft Sequence Produced by Berkeley PGA
            pga.lbl.gov
            Center Code: PGABERK.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2480: contig of 2480 bp in length
            *     2481     2580: gap of unknown length
            *     2581   100301: contig of 97721 bp in length
            *   100302   100401: gap of unknown length
            *   100402   235097: contig of 134696 bp in length.
FEATURES             Location/Qualifiers
     source          1. .235097
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="RP23-345I21"
BASE COUNT    64117 a  51797 c  51868 g  67115 t    200 others
ORIGIN
```

Alignment Scores:

Pred. No.:	8.47e-100	Length:	235097
Score:	1367.50	Matches:	252
Percent Similarity:	82.19%	Conservative:	48
Best Local Similarity:	69.04%	Mismatches:	65
Query Match:	63.40%	Indels:	1
DB:	2	Gaps:	0

US-09-208-629F-6 (1-408) x AC110877 (1-235097)

```
Qy         44 AsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaPro 63
              ||:   ::::::|||   ||||||||||||||   ||||||:::|||   |||   |||
Db     228794 AATGTTTCAGACAACTCAGCAAAGCCAACCTTAACTATTAAGAGTTTTAATGGGGGTCCC 228735

Qy         64 ProAsnSerPheGluGluProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIle 83
                 |||:::|||||||||||||||   |||   :::|||||||||||||||||||||
Db     228734 CAAAATACCTTTGAAGAATTCCCACTTTCTGACATAGAGGGCTGGACAGGAGCCACCACA 228675

Qy         84 ThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThr 103
              |||:::|||   :::|||||||||:::|||   |||   |||||||||   |||||||||
Db     228674 ACTATAAAAGCGGAGTGTCCCGAGGACAGTATTTCAACTCTCCACGTGAATAATGCTACC 228615

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
              :::|||||||||   |||||||||||||||:::::::|||||||||||||||:::|||:::
Db     228614 ATAGGATACCTGAGAAGTTCCTTAAGTACCCAAGTGATACCTGCCATCTACATCCTGCTG 228555

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArg 143
              ||||||||||||||||||||||   ||||||||||||   |||       ||||||:::
Db     228554 TTTGTGGTTGGTGTACCAGCCAACATCGTGACCCTGTGGAAACTCTCCTTAAGGACCAAA 228495

Qy        144 SerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
              ||||||      :::|||:::||||||||||||||||||||||  ||||||||||||
Db     228494 TCCATCAGTCTGGTCATCTTTCACACCAACCTGGCCATCGCAGATCTCCTTTTCTGTGTC 228435

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228434 ACACTGCCATTTAAGATCGCCTACCATCTCAATGGCAACAACTGGGTATTTGGCGAGGTC 228375

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
              ||||||   |||||||||:::|||||||||||||||||||||:::||||||:::|||
Db     228374 ACGTGCCGGATCACCACGGTTGTTTTCTACGGCAACATGTACTGCGCTATCCTGATCCTC 228315

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
              |||:::   ||||||||||||||||||      ||||||||||||||:::   |||
Db     228314 ACCTGCATGGGCATCAACCGCTACCTGGCCACGGCTCACCCTTTCACATACCAGAAGCTG 228255

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
              ||||||   ::::::::::::::   ||||||:::||||||      ||||||||||||
Db     228254 CCCAAACGCAGCTTCTCCATGCTCATGTGTGGCATGGTGTGGGTCATGGTTTTCTTATAC 228195

Qy        244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
              ||||||||||||   ||||||||||||||||||:::||||||      :::|||||||||
Db     228194 ATGCTGCCCTTTGTCATCCTGAAGCAGGAGTACCACCTCGTCCACTCCGAGATCACCACC 228135

Qy        264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
              ||||||||||||   :::   |||||||||   |||   |||:::   |||||||||:::
Db     228134 TGCCACGATGTCGTTGACGCGTGCGAGTCCCCATCATCCTTCCGATTCTACTACTTCGTC 228075

Qy        284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
              ||||||||||||||||||||||||||||||||||||||:::||||||:::||||||
Db     228074 TCCTTAGCATTCTTTGGGTTCCTCATCCCGTTTGTGATCATCATCTTCTGTTACACGACT 228015

Qy        304 IleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeu 323
              :::|||      |||   :::   |||      ||||||   |||:::||||||   |||
Db     228014 CTCATCCACAAACTTAAATCAAAGGATCGGATATGGCTGGGCTACATCAAGGCCGTCCTC 227955

Qy        324 LeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHis 343
              ||||||||||||||||||||||||||||||||||||:::||||||||||||:::||||||
Db     227954 CTCATCCTTGTGATTTTCACAATTTGCTTTGCCCCCACCAACATCATACTCGTAATCCAC 227895

Qy        344 HisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeu 363
              ||||||||||||||||||:::||||||||   |||||||||:::||||||||||||||||
Db     227894 CATGCCAACTACTACTACCACAATACCGACAGCTTGTACTTTATGTATCTTATTGCTCTG 227835

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db     227834 TGCCTGGGGAGCCTGAATAGCTGCCTAGATCCATTCCTTTACTTTGTCATGTCGAAAGTT 227775

Qy        384 ArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGluGlnGlyGln 403
              :::   |||               :::      :::   ||||||:::||||||:::
Db     227774 GTAGATC--AGCTTAATCCTTAGTCAGCAATGGCAAGACCACCTTAGAGACCAAGGAGAG 227717

Qy        404 ProSerGlnArgThr 408
              |||   :::|||
Db     227716 ATATCTGGGAAGACG 227702
```

```
RESULT 11
AR070425
LOCUS       AR070425               1124 bp    DNA     linear   PAT 18-FEB-2000
DEFINITION  Sequence 2 from patent US 5892014.
ACCESSION   AR070425
VERSION     AR070425.1  GI:7221313
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1124)
  AUTHORS   Coughlin,S.R., Ishihara,H. and Connolly,A.
  TITLE     DNA encoding a protease-activated receptor 3
  JOURNAL   Patent: US 5892014-A 2 06-APR-1999;
FEATURES             Location/Qualifiers
     source          1. .1124
                     /organism="unknown"
BASE COUNT      251 a    314 c    210 g    326 t     23 others
ORIGIN
```

Alignment Scores:

Pred. No.:	3.09e-69	Length:	1124
Score:	959.00	Matches:	185
Percent Similarity:	73.99%	Conservative:	34
Best Local Similarity:	62.50%	Mismatches:	53

```
Query Match:            44.46%      Indels:        25
DB:                     6           Gaps:          2

US-09-208-629F-6 (1-408) x AR070425 (1-1124)

Qy         44 AsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaPro 63
              |||   ::::::|||   |||||||||||||||   ||||||:::|||   |||   |||
Db         36 AATGTTTCAGACAACTCAGCAAAGCCAACCTTAACTATTAAGAGTTTTAATGGGGGTCCC 95

Qy         64 ProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIle 83
                 |||:::||||||||||||
Db         96 CAAAATACCTTTGAAGAATTCNN---------------------------------- 118

Qy         84 ThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThr 103
                                              :::   ||||||||   |||||||||
Db        119 --------------------------NNNNNTACAACTCTCCATGTGAATAATGCTACC 151

Qy        104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
              ||||||||||||   |||||||||||||||||:::|||||||||||||||:::||||||
Db        152 ATGGGATACCTGAGAAGTTCCTTAAGTACCAAAGTGATACCTGCCATCTACATCCTGGTG 211

Qy        124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArg 143
              ||||||:::|||||||||||||   ||||||||||||   |||       ||||||:::
Db        212 TTTGTGATTGGTGTACCAGCGAACATCGTGACCCTGTGGAAACTCTCCTCAAGGACCAAA 271

Qy        144 SerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
              |||||||||      :::|||:::|||||||||||||||||||||   ||||||||||||
Db        272 TCCATCTGTCTGGTCATCTTTCACACCAACCTGGCCATCGCGGATCTCCTTTTCTGTGTC 331

Qy        164 ThrLeuProPheLysIle-AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVa 183
              ||||||||||||||||||    |||||||||||||||||||||||||||||||||||||
Db        332 ACGCTGCCGTTTAAGATCNNCCTACCATCTCAATGGCAACAACTGGGTATTTGGCGAGGT 391

Qy        183 lLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer---IleLeuLe 202
              |:::||||||   |||||||||:::|||||||||||||||||||:::      |||::
Db        392 CATGTGCCGGATCACCACGGTCGTTTTCTACGGCAACATGTACTGCGCTANNNTCCTGAT 451

Qy        202 uLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGl 222
              :|||   |||:::   |||||||||||||||||      ||||||||||||||:::
Db        452 CCTCACCTGCATGGGCATCAACCGCTACCTGGCCACGGCTCACCCTTTCACATACCAGAA 511

Qy        222 yLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLe 242
               |||||||||   ::::::::::::::::   ||||||:::||||||      ||||||||
Db        512 GCTGCCCAAACGCAGCTTCTCCATGCTCATGTGTGGCATGGTGTGGGTCATGGTTTTCTT 571

Qy        242 uTyrMetLeuProPhePheIle-LeuLysGlnGluTyrTyrLeuValGlnProAspIleT 262
              ||||||||||||||||   |||   ||||||||||||:::||||||     :::||||
Db        572 ATACATGCTGCCCTTTGTCATCCNNNAAGCAGGAGTACCACCTCGTCCACTCCGAGATCA 631

Qy        262 hrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrP 282
              ||||||||||||||||   :::   |||   |||   |||   |||:::   |||||||
Db        632 CCACCTGCCACGATGTCGTCGACGCGTGCGANTCCCCATCATCCTTCCGATTCTACTACT 691

Qy        282 heIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrA 302
              ||:::||||||||||||||||||||||||||||||||||:::||||||:::||||||
Db        692 TCGTCTCCTTAGCATTCTTTGGGTTCCTCATCCCGTTTGTGATCATCATCTTCTGTTACA 751

Qy        302 laAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaS 322
                   :::|||      |||   :::   |||      ||||||   |||:::||||||
Db        752 CGACTCTCATCCACAAACTTAAATCAAAAGATCNGATATGGCTGGGCTACATCAAGGCCG 811

Qy        322 erLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
              ||||||||||||||   ||||||||||||||   ||:::
Db        812 TCCTCCTCATCCTTGTGAATTTCACCATCTGCTTCCCCCCCACC 855

RESULT 12
MMPAR2MR
LOCUS       MMPAR2MR               2713 bp    mRNA    linear   ROD 08-JUN-1995
DEFINITION  M.musculus mRNA for proteinase activated receptor 2.
ACCESSION   Z48043
VERSION     Z48043.1  GI:663020
KEYWORDS    proteinase activated receptor 2.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2713)
  AUTHORS   Nystedt,S., Larsson,A.K., Aberg,H. and Sundelin,J.
  TITLE     The mouse proteinase-activated receptor-2 cDNA and gene. Molecular
            cloning and functional expression
  JOURNAL   J. Biol. Chem. 270 (11), 5950-5955 (1995)
  MEDLINE   95197620
   PUBMED   7890726
REFERENCE   2  (bases 1 to 2713)
  AUTHORS   Nystedt,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-1995) Sverker Nystedt, Division of Molecular
            Neurobiology, The Wallenberg Laboratory, Lund University,
            Soelvegatan 33A, Lund, S-220 07, Sweden
FEATURES             Location/Qualifiers
     source          1. .2713
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="cm140pBS"
     5'UTR           1. .72
                     /citation=[1]
     CDS             73. .1272
                     /citation=[1]
                     /codon_start=1
                     /product="proteinase activated receptor 2"
                     /protein_id="CAA88097.1"
                     /db_xref="GI:663021"
                     /db_xref="MGD:MGI:101910"
                     /db_xref="SWISS-PROT:P55086"
                     /translation="MRSLSLAWLLGGITLLAASVSCSRTENLAPGRNNSKGRSLIGRL
                     ETQPPITGKGVPVEPGFSIDEFSASILTGKLTTVFLPVVYIIVFVIGLPSNGMALWIF
                     LFRTKKKHPAVIYMANLALADLLSVIWFPLKISYHLHGNNWVYGEALCKVLIGFFYGN
                     MYCSILFMTCLSVQRYWVIVNPMGHPRKKANIAVGVSLAIWLLIFLVTIPLYVMKQTI
                     YIPALNITTCHDVLPEEVLVGDMFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAM
                     DEHSEKKRQRAIRLIITVLAMYFICFAPSNLLLVVHYFLIKTQRQSHVYALYLVALCL
                     STLNSCIDPFVYYFVSKDFRDHARNALLCRSVRTVNRMQISLSSNKFSRKSGSYSSSS
                     TSVKTSY"
     intron          163^164
                     /note="genomic intron location"
                     /citation=[1]
     3'UTR           1273. .2713
                     /citation=[1]
     polyA_signal    2219. .2224
     polyA_signal    2693. .2698
BASE COUNT      632 a    676 c    633 g    772 t
ORIGIN

Alignment Scores:
Pred. No.:              4.02e-40      Length:         2713
Score:                  606.50        Matches:        137
Percent Similarity:     53.23%        Conservative:   69
Best Local Similarity:  35.40%        Mismatches:     126
Query Match:            28.12%        Indels:         55
DB:                     10            Gaps:           12

US-09-208-629F-6 (1-408) x MMPAR2MR (1-2713)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
              :::|||   :::         |||:::   ||||||      :::         |||
Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
                 ||||||               |||   :::|||             |||
Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProPheSerAlaLeuGluGlyTrpThr 79
                    ||||||                                    |||
Db        205 GAAACCCAGCCTCCA-------------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97
              |||      :::   |||:::                  |||   |||||||||
```

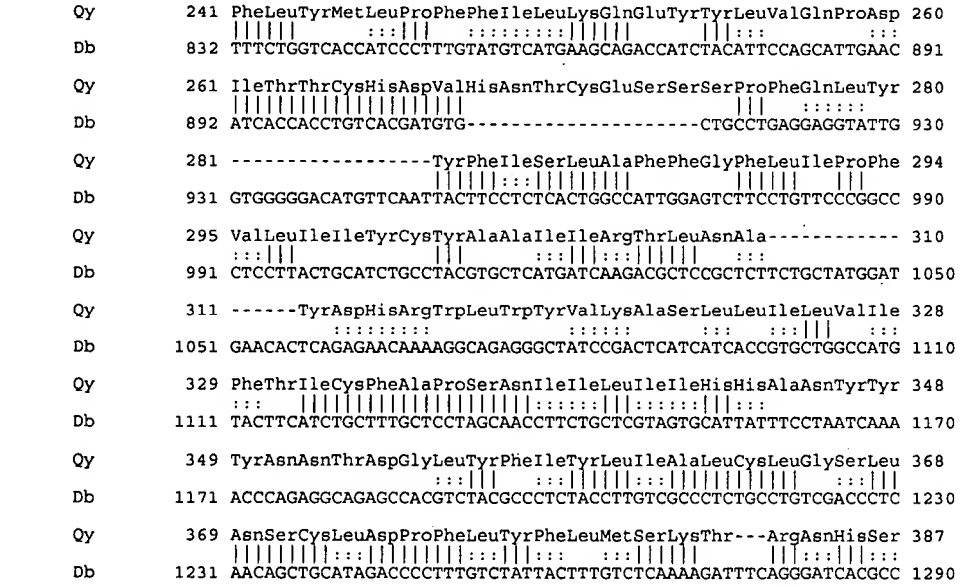
```
Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117

Db        280 ----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137

Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156

Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176

Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196

Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216

Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235

Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253

Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273

Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG------------------ 765

Qy        274 SerSerProPheGlnLeuTyr------------------TyrPheIleSerLeuAlaPhe 287

Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307

Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla------------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321

Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341

Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361

Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381

Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387

Db       1123 AAAGATTTCAGGGATCACGCC 1143
```

```
RESULT 13
AR012639
LOCUS       AR012639               2732 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 60 from patent US 5763575.
ACCESSION   AR012639
VERSION     AR012639.1  GI:3970629
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2732)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Agonist and antagonist peptides of the C140 receptor
  JOURNAL   Patent: US 5763575-A 60 09-JUN-1998;
FEATURES             Location/Qualifiers
     source          1. .2732
                     /organism="unknown"
BASE COUNT      650 a    676 c    633 g    773 t
ORIGIN

Alignment Scores:
Pred. No.:                 4.05e-40        Length:          2732
Score:                     606.50          Matches:         137
Percent Similarity:        53.23%          Conservative:    69
Best Local Similarity:     35.40%          Mismatches:      126
Query Match:               28.12%          Indels:          55
DB:                        6               Gaps:            12

US-09-208-629F-6 (1-408) x AR012639 (1-2732)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGly 41

Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59

Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProProPheSerAlaLeuGluGlyTrpThr 79

Db        205 GAAACCCAGCCTCCA---------------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97

Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117

Db        280 ----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137

Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156

Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176

Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196

Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216

Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235

Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253

Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
```

```
Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG----------------- 765

Qy        274 SerSerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhe 287
                  |||   ::::::                     ||||||:::|||||||||
Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
                  ||||||   |||   :::|||             |||      :::|||:::|||
Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla-----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321
                  |||   :::                  :::   :::          ::::::
Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
                 :::   :::|||   ::::::   ||||||||||||||||||||:::::|||:::
Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
                 :::|||:::                              :::|||   :::||||:::
Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
                 ||||||||||||   :::||||||||||||:::||||||||:::|||:::   :::|||
Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387
                 |||      |||:::|||:::
Db       1123 AAAGATTTCAGGGATCACGCC 1143

RESULT 14
AR171259
LOCUS       AR171259               2732 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 60 from patent US 6297026.
ACCESSION   AR171259
VERSION     AR171259.1  GI:17910209
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2732)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Nucleic acids encoding the C140 receptor
  JOURNAL   Patent: US 6297026-A 60 02-OCT-2001;
FEATURES             Location/Qualifiers
     source          1. .2732
                     /organism="unknown"
BASE COUNT      650 a    676 c    633 g    773 t
ORIGIN

Alignment Scores:
Pred. No.:              4.05e-40        Length:         2732
Score:                  606.50          Matches:        137
Percent Similarity:     53.23%          Conservative:   69
Best Local Similarity:  35.40%          Mismatches:     126
Query Match:            28.12%          Indels:         55
DB:                     6               Gaps:           12

US-09-208-629F-6 (1-408) x AR171259 (1-2732)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
                 :::|||   :::         |||:::   ||||||      :::          |||
Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
                 ||||||               |||   :::|||                |||
Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluProPheSerAlaLeuGluGlyTrpThr 79
                    ||||||                                          |||
Db        205 GAAACCCAGCCTCCA-----------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97
                    |||      :::   |||:::                  |||   |||||||||
Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
                                          ||||||      |||:::|||      :::||||
Db        280 -----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137
                 :::|||:::::|||||||||:::|||:::|||:::|||   :::   ||||||:::
Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
                 ||||||||||:::                  |||   |||      |||||||||:::
Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
                 ||||||   |||       :::      |||   ||||||:::|||||||||:::|||||||
Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
                 |||||||||:::||||||   ||||||:::                  ||||||||||||||
Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
                 ||||||||||||||||   :::   |||:::|||:::   ||||||      ||||||:::
Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
                 |||              |||   |||:::          |||   :::      |||:::
Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
                       :::|||      :::||||||      :::|||   :::::::::||||||
Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
                    |||:::         :::||||||||||||||||||||
Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG----------------- 765

Qy        274 SerSerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhe 287
                    |||   ::::::                     ||||||:::|||||||||
Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
                    ||||||   |||   :::|||             |||      :::|||:::|||
Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla-----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321
                    |||   :::                  :::   :::          ::::::
Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
                    :::   :::|||   ::::::   ||||||||||||||||||||:::::|||:::
Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
                    :::|||:::                              :::|||   :::||||:::
Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
                    ||||||||||||   :::||||||||||||:::||||||||:::|||:::   :::|||
Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387
                    |||      |||:::|||:::
Db       1123 AAAGATTTCAGGGATCACGCC 1143

RESULT 15
I87850
LOCUS       I87850                 2732 bp    DNA     linear   PAT 10-AUG-1998
DEFINITION  Sequence 60 from patent US 5716789.
ACCESSION   I87850
VERSION     I87850.1  GI:3407790
```

```
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2732)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Method to determine ligands, agonist and antagonist of C140
            receptor
  JOURNAL   Patent: US 5716789-A 60 10-FEB-1998;
FEATURES             Location/Qualifiers
     source          1. .2732
                     /organism="unknown"
BASE COUNT      650 a    676 c    633 g    773 t
ORIGIN

Alignment Scores:
Pred. No.:              4.05e-40        Length:         2732
Score:                  606.50          Matches:        137
Percent Similarity:     53.23%          Conservative:   69
Best Local Similarity:  35.40%          Mismatches:     126
Query Match:            28.12%          Indels:         55
DB:                     6               Gaps:           12

US-09-208-629F-6 (1-408) x I87850 (1-2732)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
              :::|||   :::         |||:::  ||||||      :::         |||
Db         85 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 144

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
              ||||||          |||  :::|||           |||
Db        145 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTA 204

Qy         60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
                  ||||||                                          |||
Db        205 GAAACCCAGCCTCCA-------------------------------------ATCACT 225

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97
              |||      :::   |||:::                  |||   |||||||||
Db        226 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 279

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
                                    ||||||      |||:::|||      :::|||
Db        280 ----------------------ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCG 315

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137
              :::|||:::::|||||||||:::|||:::|||:::|||   :::   ||||||:::
Db        316 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 375

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
                 |||||||||:::                 |||   |||      |||||||||:::
Db        376 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 435

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
              ||||||   |||      :::      |||   ||||||:::|||||||||:::||||||
Db        436 GCCGACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAAC 495

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
              |||||||||:::||||||   ||||||:::                |||||||||||||||
Db        496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATG 555

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
              ||||||||||||||   :::  |||:::|||:::  ||||||      ||||||:::
Db        556 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
              |||          |||   |||:::      |||   :::      |||:::
Db        616 CCCATG---------GGACACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 663

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
                   :::|||   :::||||||      :::|||   :::::::::||||||
Db        664 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
                 |||:::         :::||||||||||||||||||||
Db        724 ATCTACATTCCAGCATTGAACATCACCACCTGTCACGATGTG---------------- 765

Qy        274 SerSerProPheGlnLeuTyr----------------TyrPheIleSerLeuAlaPhe 287
                 |||   ::::::                    ||||||:::||||||||
Db        766 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 822

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
                 ||||||   |||   :::|||           |||      :::|||:::|||
Db        823 GGAGTCTTCCTGTTCCCGGCCCTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 882

Qy        308 LeuAsnAla----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321
              |||   :::                :::   :::          ::::::
Db        883 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 942

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
                 :::   :::|||   ::::::   ||||||||||||||||||||:::::|||:::
Db        943 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1002

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
              :::|||:::                      :::|||   :::||||||:::
Db       1003 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTC 1062

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
              ||||||||||||   :::||||||||||||:::||||||||||:::|||:::   :::|||
Db       1063 GCCCTCTGCCTGTCGACCCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1122

Qy        382 LysThr---ArgAsnHisSer 387
              |||      |||:::|||:::
Db       1123 AAAGATTTCAGGGATCACGCC 1143

RESULT 16
AR012637
LOCUS       AR012637               1475 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 1 from patent US 5763575.
ACCESSION   AR012637
VERSION     AR012637.1  GI:3970627
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1475)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Agonist and antagonist peptides of the C140 receptor
  JOURNAL   Patent: US 5763575-A 1 09-JUN-1998;
FEATURES             Location/Qualifiers
     source          1. .1475
                     /organism="unknown"
BASE COUNT      341 a    414 c    319 g    401 t
ORIGIN

Alignment Scores:
Pred. No.:              2.41e-40        Length:         1475
Score:                  605.50          Matches:        140
Percent Similarity:     52.50%          Conservative:   70
Best Local Similarity:  35.00%          Mismatches:     117
Query Match:            28.07%          Indels:         74
DB:                     6               Gaps:           13

US-09-208-629F-6 (1-408) x AR012637 (1-1475)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuLeuProThr--- 36
              :::   :::|||      |||         |||   |||||||||       :::
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla----------- 50
                   ||||||   :::          |||         :::
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 -----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
                        ||||||:::|||           |||   ||||||      |||
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399
```

```
Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db        400 TTTTCCATCGATGAGTTC------------------------------------------ 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db        418 ----------------------TCTGCGTCC------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        892 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db       1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290

RESULT 17
AR171257
LOCUS       AR171257               1475 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 1 from patent US 6297026.
ACCESSION   AR171257
VERSION     AR171257.1  GI:17910207
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1475)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Nucleic acids encoding the C140 receptor
  JOURNAL   Patent: US 6297026-A 1 02-OCT-2001;
FEATURES             Location/Qualifiers
     source          1. .1475
                     /organism="unknown"
BASE COUNT      341 a    414 c    319 g    401 t
ORIGIN

Alignment Scores:
Pred. No.:                2.41e-40        Length:          1475
Score:                    605.50          Matches:         140
Percent Similarity:       52.50%          Conservative:    70
Best Local Similarity:    35.00%          Mismatches:      117
Query Match:              28.07%          Indels:          74
DB:                       6               Gaps:            13

US-09-208-629F-6 (1-408) x AR171257 (1-1475)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla---------- 50
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 ----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db        400 TTTTCCATCGATGAGTTC------------------------------------------ 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db        418 ----------------------TCTGCGTCC------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831
```



```
Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        892 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db       1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290
```

RESULT 18
I42454
LOCUS I42454 1475 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5629174.
ACCESSION I42454
VERSION I42454.1 GI:2467949
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1475)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Recombinant C140 receptor
JOURNAL Patent: US 5629174-A 1 13-MAY-1997;
FEATURES Location/Qualifiers
source 1. .1475
/organism="unknown"
BASE COUNT 341 a 414 c 319 g 401 t
ORIGIN

Alignment Scores:

Pred. No.:	2.41e-40	Length:	1475
Score:	605.50	Matches:	140
Percent Similarity:	52.50%	Conservative:	70
Best Local Similarity:	35.00%	Mismatches:	117
Query Match:	28.07%	Indels:	74
DB:	6	Gaps:	13

US-09-208-629F-6 (1-408) x I42454 (1-1475)

```
Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla---------- 50
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 ----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
```

```
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db        400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db        418 ----------------------TCTGCGTCC--------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        892 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db       1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290
```

RESULT 19
I87848
LOCUS I87848 1475 bp DNA linear PAT 10-AUG-1998

```
DEFINITION  Sequence 1 from patent US 5716789.
ACCESSION   I87848
VERSION     I87848.1  GI:3407788
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1475)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Method to determine ligands, agonist and antagonist of C140
            receptor
  JOURNAL   Patent: US 5716789-A 1 10-FEB-1998;
FEATURES             Location/Qualifiers
     source          1. .1475
                     /organism="unknown"
BASE COUNT      341 a    414 c    319 g    401 t
ORIGIN

Alignment Scores:
Pred. No.:                2.41e-40        Length:          1475
Score:                    605.50          Matches:         140
Percent Similarity:       52.50%          Conservative:    70
Best Local Similarity:    35.00%          Mismatches:      117
Query Match:              28.07%          Indels:          74
DB:                       6               Gaps:            13

US-09-208-629F-6 (1-408) x I87848 (1-1475)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla---------- 50
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 ----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db        400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db        418 ----------------------TCTGCGTCC-------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db        892 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db       1051 GAACACTCAGAGAACAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290

RESULT 20
BC025432
LOCUS       BC025432               2772 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, coagulation factor II (thrombin) receptor-like 1,
            clone MGC:29183 IMAGE:5006769, mRNA, complete cds.
ACCESSION   BC025432
VERSION     BC025432.1  GI:19343971
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2772)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.
```

```
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 39 Row: d Column: 5
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES              Location/Qualifiers
     source           1. .2772
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /map="FVB/N-3"
                      /clone="MGC:29183 IMAGE:5006769"
                      /tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
                      old mouse. Taken by biopsy."
                      /clone_lib="NCI_CGAP_Mam2"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     CDS              99. .1298
                      /codon_start=1
                      /product="coagulation factor II (thrombin) receptor-like
                      1"
                      /protein_id="AAH25432.1"
                      /db_xref="GI:19343972"
                      /db_xref="LocusID:14063"
                      /translation="MRSLSLAWLLGGITLLAASVSCSRTENLAPGRNNSKGRSLIGRL
                      ETQPPITGKGVPVEPGFSIDEFSASILTGKLTTVFLPVVYIIVFVIGLPSNGMALWIF
                      LFRTKKKHPAVIYMANLALADLLSVIWFPLAIAYHLHGNNWVYGEALCKVLIGFFYGN
                      MYCSILFMTCLSVQRYWVIVNPMGHPRKKANIAVGVSLAIWLLIFLVTIPLYVMKQTI
                      YIPALNITTCHDVLPEEVLVGDMFNYFLSLAIGVFLFPAILTASAYVLMIKTLRSSAM
                      DEHSEKKRQRAIRLIITVLAMYFICFAPSNLLLVVHYFLIKTQRQSHVYALYLVALCL
                      STLNSCIDPFVYYFVSKDFRDHARNALLCRSVRTVNRMQISLSSNKFSRKSGSYSSSS
                      TSVKTSY"
BASE COUNT      671 a    689 c    638 g    774 t
ORIGIN

Alignment Scores:
Pred. No.:              4.97e-40        Length:         2772
Score:                  605.50          Matches:        137
Percent Similarity:     52.97%          Conservative:   68
Best Local Similarity:  35.40%          Mismatches:     127
Query Match:            28.07%          Indels:         55
DB:                     10              Gaps:           12

US-09-208-629F-6 (1-408) x BC025432 (1-2772)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
              :::|||   :::         |||:::   ||||||      :::          |||
Db        111 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTGGCGGCCTCGGTCTCCTGCAGCCGG 170

Qy         42 MetGluAsn------AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
                ||||||         |||   :::|||              |||
Db        171 ACCGAGAACCTTGCACCGGGACGCAACAACAGTAAAGGAAGAAGTCTTATTGGCAGATTG 230

Qy         60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
                  ||||||                                          |||
Db        231 GAAACCCAGCCTCCA-----------------------------------ATCACT 251

Qy         80 GlyAlaThrIleThrValLys------IleLysCysProGluGluSerAlaSerHisLeu 97
              |||      :::   |||:::                  |||   |||||||||
Db        252 GGGAAAGGGGTTCCGGTAGAACCAGGCTTTTCCATCGATGAGTTCTCTGCGTCC------ 305

Qy         98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
                                    ||||||      |||:::|||     :::|||
Db        306 ----------------------ATCCTCACCGGGAAGCTGACCACTGTCTTTCTTCCG 341

Qy        118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137
               :::|||:::::|||||||||:::|||:::|||:::|||   :::   ||||||:::
Db        342 GTCGTCTACATTATTGTGTTTGTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATC 401

Qy        138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
                    |||||||||:::               |||   |||      |||||||||:::
Db        402 TTCCTTTTCCGAACGAAGAAGAAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTG 461

Qy        157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
              ||||||   |||      :::      |||      |||||||||||||||:::||||||
Db        462 GCTGACCTCCTCTCTGTCATCTGGTTCCCCCTGGCCATTGCCTACCACCTACATGGCAAC 521

Qy        177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
              |||||||||:::||||||   ||||||:::                |||||||||||||||
Db        522 AACTGGGTCTATGGGGAGGCCCTGTGCAAGGTGCTCATTGGCTTTTTCTATGGCAACATG 581

Qy        197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
              ||||||||||||||   :::  |||:::|||:::   ||||||      ||||||:::
Db        582 TATTGCTCCATCCTCTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 641

Qy        217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
              |||            |||   |||:::         |||   :::      |||:::
Db        642 CCCATG---------GGGCACCCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCC 689

Qy        236 ------ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
                   :::|||      :::||||||      :::|||   :::::::::||||||
Db        690 TTGGCAATCTGGCTCCTGATTTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 749

Qy        254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
                 |||:::         :::||||||||||||||||||||
Db        750 ATCTACATTCCAGCCTTGAACATCACCACCTGCCACGATGTG----------------- 791

Qy        274 SerSerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhe 287
                    |||   ::::::                  ||||||:::|||||||||
Db        792 ---CTGCCTGAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATT 848

Qy        288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
                    ||||||   |||   :::|||          |||      :::|||:::|||
Db        849 GGAGTCTTCCTGTTCCCGGCCATCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACG 908

Qy        308 LeuAsnAla----------------TyrAspHisArgTrpLeuTrpTyrValLysAla 321
              |||   :::                    :::   :::          ::::::
Db        909 CTCCGCTCTTCTGCTATGGATGAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTC 968

Qy        322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
                 :::   :::|||   ::::::   |||||||||||||||||||:::::|||:::
Db        969 ATCATCACCGTGCTGGCCATGTACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTA 1028

Qy        342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
              :::|||:::                    :::|||   :::||||||:::
Db       1029 GTGCATTATTTCCTAATCAAAACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTCGTC 1088

Qy        362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
              ||||||||||||   :::|||:|||||||:::|||||||:::|||:::   :::|||
Db       1089 GCCCTCTGCCTGTCGACTCTCAACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCA 1148

Qy        382 LysThr---ArgAsnHisSer 387
              |||      |||:::|||:::
Db       1149 AAAGATTTCAGGGATCACGCC 1169

RESULT 21
MMPROACR2
LOCUS       MMPROACR2               1477 bp    DNA     linear   ROD 30-JAN-1995
DEFINITION  M.musculus exon 2 for proteinase activated receptor 2.
ACCESSION   Z35158
VERSION     Z35158.1  GI:559918
KEYWORDS    proteinase activated receptor; proteinase activated receptor 2.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1477)
  AUTHORS   Nystedt,S., Emilsson,K., Wahlestedt,C. and Sundelin,J.
  TITLE     Molecular cloning of a potential proteinase activated receptor
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 91 (20), 9208-9212 (1994)
  MEDLINE   95023880
   PUBMED   7937743
REFERENCE   2  (bases 1 to 1477)
  AUTHORS   Nystedt,S.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (11-JUL-1994) Sverker Nystedt, Division of Neurobiology,
            The Wallenberg, Laboratory, Lund University, Soelvegatan 33A, Lund,
            S-220 07, Sweden
REFERENCE   3  (bases 1 to 1477)
  AUTHORS   Nystedt,S., Larsson,A.K., Aaberg,H. and Sundelin,J.
  TITLE     The mouse proteinase activated receptor 2 cDNA and gene-molecular
            cloning and functional expression
  JOURNAL   J. Biol. Chem. 170 (1995) In press
FEATURES             Location/Qualifiers
     source          1. .1477
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /germline
     gene            1. .1477
                     /gene="PAR-2"
     intron          <1. .310
                     /gene="PAR-2"
                     /citation=[3]
     exon            311. .>1477
                     /gene="PAR-2"
                     /citation=[3]
                     /number=1
   3'UTR             1420. .>1477
                     /gene="PAR-2"
BASE COUNT      341 a    413 c    321 g    402 t
ORIGIN

Alignment Scores:
Pred. No.:              3.5e-40       Length:          1477
Score:                  603.50        Matches:         140
Percent Similarity:     52.25%        Conservative:    69
Best Local Similarity:  35.00%        Mismatches:      118
Query Match:            27.98%        Indels:          74
DB:                     10            Gaps:            13

US-09-208-629F-6 (1-408) x MMPROACR2 (1-1477)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
              :::   :::|||     |||       |||   |||||||||        :::
Db        221 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla----------- 50
                  ||||||   :::          |||         :::
Db        281 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 340

Qy         51 -----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
                       ||||||:::|||          |||   ||||||      |||
Db        341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 399

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
              |||   :::||||||
Db        400 TTTTCCATCGATGAGTTC---------------------------------------- 417

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
                                    |||||||||
Db        418 ----------------------TCTGCGTCC------------------------- 426

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
                  ||||||      |||:::|||      :::|||   :::|||::::::||||||
Db        427 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 483

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
              |||:::|||:::|||:::|||   :::   ||||||:::      |||||||||:::
Db        484 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 543

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
                     |||   |||      |||||||||:::||||||   |||      :::
Db        544 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 603

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
                  |||   ||||||:::|||||||||:::|||||||||||||||:::||||||
Db        604 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 663

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
              ||||||:::                |||||||||||||||||||||||||||||   :::
Db        664 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 723

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
                 |||:::|||:::   ||||||      ||||||:::|||          |||
Db        724 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 774

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
              |||:::         |||   :::      |||:::         :::|||      :::
Db        775 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 831

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||      :::|||   :::::::::||||||      |||:::         :::
Db        832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 891

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||                        |||   ::::::
Db        892 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 930

Qy        281 -----------------TyrPheIleSerLeuAlaPheGlyPheLeuIleProPhe 294
                              ||||||:::|||||||||         ||||||   |||
Db        931 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 990

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
              :::|||            |||      :::|||:::||||||   :::
Db        991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 1050

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
                    :::   :::          ::::::      :::   :::|||   :::
Db       1051 GAACACTCAGAGAAGAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 1110

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
              :::   ||||||||||||||||||||:::::|||::::::|||:::
Db       1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 1170

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
                                  :::|||   :::|||||||:::||||||||||||   :::|||
Db       1171 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 1230

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
              |||||||||:::|||||||||:::|||:::   :::||||||      |||:::|||:::
Db       1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 1290

RESULT 22
AC110876/c
LOCUS       AC110876              199227 bp    DNA     linear   HTG 16-FEB-2002
DEFINITION  Mus musculus clone RP23-176N6, WORKING DRAFT SEQUENCE, 8 unordered
            pieces.
ACCESSION   AC110876
VERSION     AC110876.1  GI:18693418
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 199227)
  AUTHORS   Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and Cheng,J.F.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 199227)
  AUTHORS   Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and
            Cheng,J.-F.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2002) Berkeley PGA, Lawrence Berkeley National
            Laboratory, 1 Cyclotron rd., Berkeley, CA 94720, United States of
            America
COMMENT     Draft Sequence Produced by Berkeley PGA
            pga.lbl.gov
            Center Code: PGABERK.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
```

```
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     2263: contig of 2263 bp in length
            *    2264     2363: gap of unknown length
            *    2364     5169: contig of 2806 bp in length
            *    5170     5269: gap of unknown length
            *    5270    13742: contig of 8473 bp in length
            *   13743    13842: gap of unknown length
            *   13843    21596: contig of 7754 bp in length
            *   21597    21696: gap of unknown length
            *   21697    37430: contig of 15734 bp in length
            *   37431    37530: gap of unknown length
            *   37531    77278: contig of 39748 bp in length
            *   77279    77378: gap of unknown length
            *   77379   129890: contig of 52512 bp in length
            *  129891   129990: gap of unknown length
            *  129991   199227: contig of 69237 bp in length.
FEATURES             Location/Qualifiers
     source          1. .199227
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="RP23-176N6"
BASE COUNT    55269 a  45000 c  45562 g  52676 t    720 others
ORIGIN

Alignment Scores:
Pred. No.:              9.82e-38        Length:         199227
Score:                  603.50          Matches:        140
Percent Similarity:     52.25%          Conservative:   69
Best Local Similarity:  35.00%          Mismatches:     118
Query Match:            27.98%          Indels:         74
DB:                     2               Gaps:           13

US-09-208-629F-6 (1-408) x AC110876 (1-199227)

Qy         18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
              :::  :::|||     |||     |||  |||||||||       :::
Db     174666 GTTCATCTGAAATGTTCCATTTAAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 174607

Qy         37 ------PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla---------- 50
                  ||||||   :::           |||         :::
Db     174606 TAATGATTCTGCTCCGCTTTCTTTGTACAGGACGCAACAACAGTAAAGGAAGAAGTCTTA 174547

Qy      .  51 -----------LysProThrLeuProIle------LysThrPheArgGlyAlaProPro 64
                        ||||||:::|||           |||   ||||||      |||
Db     174546 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGGTTCCGGTA-GAACCAGGC 174488

Qy         65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
                 |||   :::||||||
Db     174487 TTTTCCATCGATGAGTTC----------------------------------------- 174470

Qy         85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
                                 |||||||||
Db     174469 ----------------------TCTGCGTCC------------------------- 174461

Qy        105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
                 ||||||      |||:::|||      :::|||   :::|||:::::||||||
Db     174460 ---ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCCGGTCGTCTACATTATTGTGTTT 174404

Qy        125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
               |||:::|||:::|||:::|||   :::   ||||||:::      ||||||||||:::
Db     174403 GTGATTGGTTTGCCCAGTAATGGCATGGCCCTCTGGATCTTCCTTTTCCGAACGAAGAAG 174344

Qy        145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
                        |||   |||      |||||||||:::||||||   |||      :::
Db     174343 AAACACCCCGCCGTGATTTACATGGCCAACCTGGCCTTGGCCGACCTCCTCTCTGTCATC 174284

Qy        164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
                 |||   ||||||:::|||||||||:::|||||||||||||||:::||||||
Db     174283 TGGTTCCCCCTGAAGATCTCCTACCACCTACATGGCAACAACTGGGTCTACGGGGAGGCC 174224

Qy        184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
              ||||||:::                 |||||||||||||||||||||||||||||   :::
Db     174223 CTGTGCAAGGTGCTCATTGGCTTTTTCTATGGTAACATGTATTGCTCCATCCTCTTCATG 174164

Qy        204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
                 |||:::|||:::   ||||||      ||||||:::|||                |||
Db     174163 ACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG---------GGACAC 174113

Qy        224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu---------ValTrpAlaThrVal 240
              |||:::         |||   :::      |||:::          :::|||      :::
Db     174112 CCCAGGAAGAAG---GCAAACATCGCCGTTGGCGTCTCCTTGGCAATCTGGCTCCTGATT 174056

Qy        241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
              ||||||      :::|||   :::::::::||||||      |||:::         :::
Db     174055 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTGAAC 173996

Qy        261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
              ||||||||||||||||||||                        |||   ::::::
Db     173995 ATCACCACCTGTCACGATGTG--------------------CTGCCTGAGGAGGTATTG 173957

Qy        281 ------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
                               ||||||:::|||||||||        ||||||   |||
Db     173956 GTGGGGGACATGTTCAATTACTTCCTCTCACTGGCCATTGGAGTCTTCCTGTTCCCGGCC 173897

Qy        295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------------ 310
              :::|||             |||      :::|||:::||||||   :::
Db     173896 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGAT 173837

Qy        311 ------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
                    :::   :::          ::::::      :::   :::|||   :::
Db     173836 GAACACTCAGAGAAGAAAAAGGCAGAGGGCTATCCGACTCATCATCACCGTGCTGGCCATG 173777

Qy        329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyr 348
              :::   ||||||||||||||||||||::::::|||::::::|||:::
Db     173776 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATCAAA 173717

Qy        349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
                           :::|||   :::||||||:::||||||||||||   :::|||
Db     173716 ACCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTTGTCGCCCTCTGCCTGTCGACCCTC 173657

Qy        369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
              ||||||||:::|||||||||:::|||:::   :::||||||      |||:::|||:::
Db     173656 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCACGCC 173597

RESULT 23
XLU09632
LOCUS       XLU09632               1680 bp    mRNA    linear   VRT 28-MAY-1994
DEFINITION  Xenopus laevis thrombin receptor mRNA, complete cds.
ACCESSION   U09632
VERSION     U09632.1  GI:495197
KEYWORDS    .
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 420)
  AUTHORS   Gerszten,R.E., Chen,J., Ishii,M., Ishii,K., Wang,L., Nanevicz,T.,
            Turck,C.W., Vu,T.K. and Coughlin,S.R.
  TITLE     Specificity of the thrombin receptor for agonist peptide is defined
            by its extracellular surface
  JOURNAL   Nature 368 (6472), 648-651 (1994)
  MEDLINE   94195429
   PUBMED   8145852
REFERENCE   2  (bases 1 to 1680)
  AUTHORS   Coughlin,S.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-1994) Shaun R. Coughlin, CVRI, University of
            California, San Francisco, 505 Parnassus Avenue, Room HSW 831, San
            Francisco, CA 94143, USA
FEATURES             Location/Qualifiers
     source          1. .1680
```

```
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
     CDS             197. .1459
                     /codon_start=1
                     /product="thrombin receptor"
                     /protein_id="AAA18498.1"
                     /db_xref="GI:495198"
                     /translation="MMELRVLLLLLLLTLLGAMGSLCLANSDTQAKGAHSNNMTIKTF
                     RIFDDSESEFEEIPWDELDESGEGSGDQAPVSRSARKPIRRNITKEAEQYLSSQWLTK
                     FVPSLYTVVFIVGLPLNLLAIIIFLFKMKVRKPAVVYMLNLAIADVFFVSVLPFKIAY
                     HLSGNDWLFGPGMCRIVTAIFYCNMYCSVLLIASISVDRFLAVVYPMHSLSWRTMSRA
                     YMACSFIWLISIASTIPLLVTEQTQKIPRLDITTCHDVLDLKDLKDFYIYYFSSFCLL
                     FFFVPFIITTICYIGIIRSLSSSSIENSCKKTRALFLAVVVLCVFIICFGPTNVLFLT
                     HYLQEANEFLYFAYILSACVGSVSCCLDPLIYYYASSQCQRYLYSLLCCRKVSEPGSS
                     TGQLMSTAMKNDNCSTNAKSSIYKKLLA"
BASE COUNT      427 a    366 c    401 g    486 t
ORIGIN

Alignment Scores:
Pred. No.:              8.59e-40        Length:         1680
Score:                  599.50          Matches:        128
Percent Similarity:     52.97%          Conservative:   68
Best Local Similarity:  34.59%          Mismatches:     149
Query Match:            27.79%          Indels:         25
DB:                     5               Gaps:           8

US-09-208-629F-6 (1-408) x XLU09632 (1-1680)

Qy         30 LeuLeuLeuLeuLeuProThr--------------------PheCysGlnSerGlyMet 42
              ||||||||||||||||   |||                          |||   :::
Db        218 CTGCTGCTGCTGCTATTGACCCTGCTGGGAGCGATGGGATCTCTCTGTCTGGCCAACTCT 277

Qy         43 GluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg------ 60
              :::                      :::   ||||||||||||||||
Db        278 GATACCCAAGCAAAGGGTGCTCATTCTAACAATATGACCATAAAAACTTTCAGAATATTT 337

Qy         61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeu---GluGlyTrpThr 79
                    :::   |||||||||   |||:::      |||   |||
Db        338 GACGACAGTGAAAGTGAGTTTGAAGAGATTCCATGGGATGAACTCGATGAATCAGGAGAA 397

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||:::                :::                              ::::::
Db        398 GGTTCAGGAGACCAGGCCCCGGTGTCCCGTAGTGCAAGGAAGCCAATCCGGAGGAACATC 457

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                         ||||||:::|||       ||||||   :::|||::::::
Db        458 ACAAAGGAGGCGGAGCAATATCTCTCCAGTCAGTGGCTGACCAAGTTTGTACCATCCCTG 517

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   :::||||||:::|||||||:::|||   |||   :::   :::   :::
Db        518 TACACTGTGGTGTTTATTGTGGGCCTGCCTCTGAATCTGTTGGCAATCATCATATTCCTG 577

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||:::   :::                |||   |||      ||||||||||||||||||
Db        578 TTCAAGATGAAAGTCAGAAAACCGGCAGTGGTCTATATGTTGAACCTGGCCATCGCCGAC 637

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
                  |||         |||||||||||||||||||||||||:::|||||||:::|||
Db        638 GTGTTCTTTGTTAGTGTGTTGCCTTTCAAAATAGCCTACCATCTGTCCGGAAACGACTGG 697

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              :::||||||      :::||||||      |||   |||||||||   ||||||||||||
Db        698 CTGTTTGGGCCTGGAATGTGCCGCATTGTCACTGCCATCTTCTACTGCAACATGTACTGC 757

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||:::||||||:::|||   ||||||:::::::|||:::||||||:::|||:::|||
Db        758 TCTGTTCTGCTCATCGCCAGCATCAGTGTGGACAGGTTCTTGGCAGTGGTTTATCCCATG 817

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
                                  :::   |||   :::   |||      :::|||
Db        818 CACTCCCTTTCTTGGCGCACAATGAGCCGTGCCTATATGGCCTGTAGTTTCATTTGGCTG 877

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
                                  :::|||      :::   :::|||         :::   :::
Db        878 ATATCCATAGCGAGCACTATACCACTTCTTGTGACTGAGCAAACTCAGAAGATTCCTAGG 937

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
                 ||||||||||||||||||||||   :::      :::            |||
Db        938 CTGGATATTACTACTTGCCATGATGTTTTAGAT---TTGAAAGACCTCAAAGACTTTTAT 994

Qy        279 LeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIle 298
              :::|||||||||   |||                |||   :::||||||::::::
Db        995 ATCTACTATTTCTCTTCTTTCTGCTTGCTGTTCTTCTTCGTGCCATTCATTATCACCACT 1054

Qy        299 TyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp----------------- 312
                 ||||||     |||||||||:::|||::::::
Db       1055 ATCTGCTACATAGGGATCATCAGAAGCTTAAGTTCCTCCAGCATCGAGAACAGTTGCAAG 1114

Qy        313 ---HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIle 331
                    |||   |||:::           :::::::::::|||   :::|||   |||
Db       1115 AAGACAAGAGCCCTGTTC---------TTGGCGGTGGTTGTGCTCTGTGTTTTTATCATC 1165

Qy        332 CysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsn 351
              ||||||   |||:::|||::::::   :::   |||         |||
Db       1166 TGTTTTGGGCCAACCAATGTCCTCTTTCTGACCCAT---------TACTTGCAAGAAGCC 1216

Qy        352 ThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
                 :::   |||||||||   |||:::::::::   |||:::||||||::::::   |||
Db       1217 AATGAATTTCTATATTTTGCCTACATTCTGAGTGCCTGTGTTGGCAGCGTCAGCTGTTGC 1276

Qy        372 LeuAspProPheLeuTyrPheLeuMetSer 381
              |||||||||   :::|||:::      |||
Db       1277 CTCGATCCCCTTATTTACTACTATGCGTCA 1306

RESULT 24
RNU61373
LOCUS       RNU61373                1428 bp    mRNA    linear   ROD 08-AUG-1996
DEFINITION  Rattus norvegicus proteinase-activated receptor-2 mRNA, complete
            cds.
ACCESSION   U61373
VERSION     U61373.1  GI:1480687
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1428)
  AUTHORS   Saifeddine,M., al-Ani,B., Cheng,C.H., Wang,L. and Hollenberg,M.D.
  TITLE     Rat proteinase-activated receptor-2 (PAR-2): cDNA sequence and
            activity of receptor-derived peptides in gastric and vascular
            tissue
  JOURNAL   Br. J. Pharmacol. 118 (3), 521-530 (1996)
  MEDLINE   96358009
   PUBMED   8762073
REFERENCE   2  (bases 1 to 1428)
  AUTHORS   Al-Ani,B., Hollenberg,M.D., Saifeddine,M., Cheng,C.-H. and Wang,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-1996) Pharmacology, Univ. of Calgary, 3330
            Hospital Dr NW, Calgary, AB T2N 4N1, Canada
FEATURES             Location/Qualifiers
     source          1. .1428
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="intestinal and kidney"
     CDS             1. .1194
                     /note="PAR-2"
                     /codon_start=1
                     /product="proteinase-activated receptor-2"
                     /protein_id="AAC52703.1"
                     /db_xref="GI:1480688"
                     /translation="MRSLSLAWLLGGITLLAASASCNRTVNAPGPNSKGRSLIGRLDT
                     PPPITGKGAPVEPGFSVDEFSASVLTGKLTTVFLPVIYIIVFVIGLPSNGMALWVFFF
                     RTKKKHPAVIYMANLALADLLSVIWFPLKISYHLHGNDWTYGDALCKVLIGFFYGNMY
                     CSILFMTCLSVQRYWVIVNPMGHSRKRANIAVGVSLAIWLLIFLVTIPLYVMRQTIYI
```

```
                 PALNITTCHDVLPEEVLVGDMFSYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMDE
                 HSEKKRRRAIRLIITVLSMYFICFAPSNVLLVVHYFLIKSQRQSHVYALYLVALCLST
                 LNSCIDPFVYYFVSKDFRDQARNALLCRSVRTVKRMQISLTSNKFSRKSSSYSSSSTS
                 VKTSY"
BASE COUNT      282 a    413 c    358 g    375 t
ORIGIN

Alignment Scores:
Pred. No.:                1.66e-39        Length:         1428
Score:                    595.00          Matches:        134
Percent Similarity:       51.12%          Conservative:   71
Best Local Similarity:    33.42%          Mismatches:     128
Query Match:              27.58%          Indels:         68
DB:                       10              Gaps:           12

US-09-208-629F-6 (1-408) x RNU61373 (1-1428)

Qy         22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
Db         13 AGCCTGGCGTGGCTGCTGGGAGGTATCACCCTTCTG-----------GCGGCCTCGGCC 60

Qy         42 MetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGly 61
Db         61 TCCTGCAACCGGACCGTGAATGCACCGGGACCCAACAGTAAAGGGAGAAGTCTGATTGGC 120

Qy         62 ---------AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrp 78
Db        121 AGATTGGACACGCCGCCT-----------------CCCATCACTGGG------------ 150

Qy         79 ThrGlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHis 98
Db        151 AAAGGGGCTCCAGTTGAACCAGGCTTTTCCGTTGATGAATTCTCTGCATCC--------- 201

Qy         99 ValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAla 118
Db        202 --------------------GTCCTCACCGGGAAGCTGACCACCGTCTTTCTCCCGGTC 240

Qy        119 IleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeu 138
Db        241 ATCTACATCATTGTCTTTGTAATTGGTTTGCCCAGTAATGGTATGGCCCTCTGGGTCTTC 300

Qy        139 PhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAla 157
Db        301 TTCTTCCGAACGAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAACCTGGCCTTGGCA 360

Qy        158 AspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 177
Db        361 GACCTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTCCATGGCAACGAC 420

Qy        178 TrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyr 197
Db        421 TGGACCTATGGGGATGCGCTCTGCAAGGTGCTCATTGGCTTTTTCTACGGCAATATGTAC 480

Qy        198 CysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisPro 217
Db        481 TGCTCCATCCTTTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCC 540

Qy        218 PheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu------ 235
Db        541 ATGGGACAC-----------TCCAGGAAGAGGGCCAACATCGCTGTTGGCGTCTCCCTG 588

Qy        236 ---ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyr 254
Db        589 GCCATCTGGCTCCTGATTTTTCTGGTCACCATCCCTCTGTACGTCATGAGGCAGACCATC 648

Qy        255 TyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSer 274
Db        649 TACATTCCAGCCTTGAACATCACCACCTGTCACGACGTG-------------------- 687

Qy        275 SerProPheGlnLeuTyr-----------------TyrPheIleSerLeuAlaPhePhe 288
Db        688 CTGCCCGAGGAGGTCCTGGTGGGGGACATGTTCAGTTACTTCCTCTCCCTGGCCATTGGA 747

Qy        289 GlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeu 308
Db        748 GTCTTTCTGTTCCCAGCCCTCCTTACTGCGTCTGCCTACGTGCTCATGATCAAAACGCTC 807

Qy        309 AsnAla-----------------TyrAspHisArgTrpLeuTrpTyrValLysAlaSer 322
Db        808 CGCTCTTCCGCCATGGACGAGCACTCGGAGAAGAAAAGGCGGAGGGCTATCCGCCTCATC 867

Qy        323 LeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIle 342
Db        868 ATCACGGTGCTGTCCATGTACTTCATCTGCTTCGCTCCCAGCAACGTGCTGCTCGTCGTG 927

Qy        343 HisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAla 362
Db        928 CATTATTTCCTCATCAAAAGCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTCGTCGCC 987

Qy        363 LeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLys 382
Db        988 CTCTGCCTGTCCACCCTCAACAGCTGCATAGACCCCTTTGTCTACTACTTTGTTTCGAAA 1047

Qy        383 ThrArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGluGlnGly 402
Db       1048 ----------------------------------------GATTTCAGGGACCAGGCC 1065

Qy        403 Gln 403
Db       1066 AGA 1068
```

```
RESULT 25
BC012453
LOCUS       BC012453               2813 bp    mRNA    linear   PRI 20-AUG-2001
DEFINITION  Homo sapiens, clone MGC:9549 IMAGE:3857382, mRNA, complete cds.
ACCESSION   BC012453
VERSION     BC012453.1  GI:15214649
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2813)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalon@bcm.tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,
            Muzny,D.M., Gibbs,R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 21 Row: b Column: 20
            This clone was selected for full length sequencing because it
            passed the following selection criteria: GenomeScan gene
            prediction, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .2813
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:9549 IMAGE:3857382"
                     /tissue_type="Ovary, adenocarcinoma"
```

```
                     /clone_lib="NIH_MGC_66"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             93. .1286
                     /codon_start=1
                     /product="Unknown (protein for MGC:9549)"
                     /protein_id="AAH12453.1"
                     /db_xref="GI:15214650"
                     /translation="MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG
                     TSHVTGKGVTVETVFSVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLF
                     RTKKKHPAVIYMANLALADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMY
                     CSILFMTCLSVQRYWVIVNPMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFI
                     PALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDE
                     NSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLST
                     LNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTT
                     VKTSY"
BASE COUNT      757 a    614 c    604 g    838 t
ORIGIN

Alignment Scores:
Pred. No.:              5.76e-39        Length:         2813
Score:                  592.50          Matches:        142
Percent Similarity:     50.85%          Conservative:   67
Best Local Similarity:  34.55%          Mismatches:     145
Query Match:            27.47%          Indels:         57
DB:                     9               Gaps:           14

US-09-208-629F-6 (1-408) x BC012453 (1-2813)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
              |||   |||:::|||      :::           :::   |||   |||      |||
Db         84 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 143

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
              |||   |||                  |||   ||||||                ||||||
Db        144 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 182

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
                 :::|||           |||           |||
Db        183 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA-------------- 227

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
                        |||   :::   |||      |||   |||:::
Db        228 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 269

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                     |||   |||||||||                                  |||
Db        270 TCTGTGGATGAGTTTTCTGCATCT-----------------------------GTCCTC 299

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
              |||      |||:::|||      :::|||   :::|||   :::|||||||||||||||
Db        300 ACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGT 359

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
              :::|||:::|||   :::   ||||||:::      |||||||||:::
Db        360 TTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 419

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
              |||   |||      |||||||||:::||||||   |||      :::   |||
Db        420 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 479

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
              ||||||||||||||:::::|||||||||||:::::||||||   ||||||
Db        480 TTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 539

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
                        |||||||||||||||||||||||||||||   :::   |||:::
Db        540 GTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 599

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
              |||:::   ||||||      ||||||:::|||      :::               |||
Db        600 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 656

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
                    |||:::                :::|||      :::   |||      :::|||
Db        657 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCT 716

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
                 :::::::::||||||      ::::::         :::||||||||||||||||||
Db        717 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 776

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr---------------- 280
              |||                        |||   ||||||
Db        777 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 815

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
              ||||||:::|||||||||          ||||||   |||      |||
Db        816 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 875

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla----------------TyrAspHisArg 314
              |||      :::||||||   |||   :::                        :::   :::
Db        876 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 935

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
                       :::|||      :::   :::|||   ::::::   |||||||||
Db        936 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACT 995

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
              |||||||||::::::|||::::::|||          |||:::          :::   |||
Db        996 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1046

Qy        355 --------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
                      :::|||   :::|||::::::||||||||||||   :::||||||||||||
Db       1047 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1106

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
              :::|||||||||:::|||:::   :::|||      |||:::|||:::      |||
Db       1107 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1166

Qy        390 TyrLeuThrLys******AsnAspLeuArgGlu 400
                 |||   :::           :::::::::
Db       1167 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1199

RESULT 26
AC128225/c
LOCUS       AC128225               88955 bp    DNA     linear   HTG 31-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-498A19, *** SEQUENCING IN PROGRESS
            ***, 44 unordered pieces.
ACCESSION   AC128225
VERSION     AC128225.1  GI:21908835
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 88955)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
```

```
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 88955)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 88955)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GZLT
                Center clone name: CH230-498A19
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 35763 bases at least Q40
                Consensus quality: 40497 bases at least Q30
                Consensus quality: 44674 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 44 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1023: contig of 1023 bp in length
            *     1024     1123: gap of unknown length
            *     1124     2158: contig of 1035 bp in length
            *     2159     2258: gap of unknown length
            *     2259     3396: contig of 1138 bp in length
            *     3397     3496: gap of unknown length
            *     3497     4759: contig of 1263 bp in length
            *     4760     4859: gap of unknown length
            *     4860     6321: contig of 1462 bp in length
            *     6322     6421: gap of unknown length
            *     6422     7471: contig of 1050 bp in length
            *     7472     7571: gap of unknown length
            *     7572     8781: contig of 1210 bp in length
            *     8782     8881: gap of unknown length
            *     8882    10238: contig of 1357 bp in length
            *    10239    10338: gap of unknown length
            *    10339    11514: contig of 1176 bp in length
            *    11515    11614: gap of unknown length
            *    11615    13207: contig of 1593 bp in length
            *    13208    13307: gap of unknown length
            *    13308    14714: contig of 1407 bp in length
            *    14715    14814: gap of unknown length
            *    14815    16170: contig of 1356 bp in length
            *    16171    16270: gap of unknown length
            *    16271    17526: contig of 1256 bp in length
            *    17527    17626: gap of unknown length
            *    17627    19027: contig of 1401 bp in length
            *    19028    19127: gap of unknown length
            *    19128    20258: contig of 1131 bp in length
            *    20259    20358: gap of unknown length
            *    20359    21562: contig of 1204 bp in length
            *    21563    21662: gap of unknown length
            *    21663    23425: contig of 1763 bp in length
            *    23426    23525: gap of unknown length
            *    23526    24837: contig of 1312 bp in length
            *    24838    24937: gap of unknown length
            *    24938    25946: contig of 1009 bp in length
            *    25947    26046: gap of unknown length
            *    26047    27583: contig of 1537 bp in length
            *    27584    27683: gap of unknown length
            *    27684    29343: contig of 1660 bp in length
            *    29344    29443: gap of unknown length
            *    29444    30864: contig of 1421 bp in length
            *    30865    30964: gap of unknown length
            *    30965    32341: contig of 1377 bp in length
            *    32342    32441: gap of unknown length
            *    32442    34492: contig of 2051 bp in length
            *    34493    34592: gap of unknown length
            *    34593    37027: contig of 2435 bp in length
            *    37028    37127: gap of unknown length
            *    37128    40023: contig of 2896 bp in length
            *    40024    40123: gap of unknown length
            *    40124    42440: contig of 2317 bp in length
            *    42441    42540: gap of unknown length
            *    42541    45193: contig of 2653 bp in length
            *    45194    45293: gap of unknown length
            *    45294    47982: contig of 2689 bp in length
            *    47983    48082: gap of unknown length
            *    48083    49458: contig of 1376 bp in length
            *    49459    49558: gap of unknown length
            *    49559    51558: contig of 2000 bp in length
            *    51559    51658: gap of unknown length
            *    51659    53031: contig of 1373 bp in length
            *    53032    53131: gap of unknown length
            *    53132    55703: contig of 2572 bp in length
            *    55704    55803: gap of unknown length
            *    55804    57502: contig of 1699 bp in length
            *    57503    57602: gap of unknown length
            *    57603    59938: contig of 2336 bp in length
            *    59939    60038: gap of unknown length
            *    60039    62314: contig of 2276 bp in length
            *    62315    62414: gap of unknown length
            *    62415    65536: contig of 3122 bp in length
            *    65537    65636: gap of unknown length
            *    65637    68901: contig of 3265 bp in length
            *    68902    69001: gap of unknown length
            *    69002    71414: contig of 2413 bp in length
            *    71415    71514: gap of unknown length
            *    71515    74439: contig of 2925 bp in length
            *    74440    74539: gap of unknown length
            *    74540    78489: contig of 3950 bp in length
            *    78490    78589: gap of unknown length
            *    78590    82199: contig of 3610 bp in length
            *    82200    82299: gap of unknown length
            *    82300    85743: contig of 3444 bp in length
```

```
           *    85744    85843: gap of unknown length
           *    85844    88955: contig of 3112 bp in length.
FEATURES             Location/Qualifiers
     source          1. .88955
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-498A19"
BASE COUNT    22906 a  20029 c  19951 g  21690 t   4379 others
ORIGIN

Alignment Scores:
Pred. No.:              3.05e-37        Length:         88955
Score:                  592.50          Matches:        121
Percent Similarity:     55.19%          Conservative:   65
Best Local Similarity:  35.91%          Mismatches:     106
Query Match:            27.47%          Indels:         45
DB:                     2               Gaps:           7

US-09-208-629F-6 (1-408) x AC128225 (1-88955)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db      33842 GGGGCTCCAGTTGAACCAGGCTTTTCCGTTGATGAATTCTCTGCATCC------------ 33795

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
Db      33794 -----------------GTCCTCACCGGGAAGCTGACCACCGTCTTTCTCCCGGTCATC 33753

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db      33752 TACATCATTGTCTTTGTAATTGGTTTGCCCAGTAATGGTATGGCCCTCTGGGTCTTCTTC 33693

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db      33692 TTCCGAACGAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAACCTGGCCTTGGCAGAC 33633

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db      33632 CTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTCCATGGCAACGACTGG 33573

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db      33572 ACCTATGGGGATGCGCTCTGCAAGGTGCTCATTGGCTTTTTCTACGGCAATATGTACTGC 33513

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db      33512 TCCATCCTTTTCATGAACTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG 33453

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db      33452 GGACACTCCAGGAAGAGGGGCAACATCGCTGTTGGCGTCTCCCTGGGC---ATCTGGCTC 33396

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db      33395 CTGATTTTTCTGGTCACCATCCCTCTGTACGTCATGAGGCAGAACATCTACATTCCAGCC 33336

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db      33335 TTGAACATCACCACCTGTCACGACGTG--------------------CTGCCCGAGGAG 33297

Qy        279 LeuTyr------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db      33296 GTCCTGGTGGGGGACATGTTCAGTTACTTCCTCTCCCTGGCCATTGGAGTCTTTCTGTTC 33237

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
Db      33236 CCAGCCCTCCTTACTGCGTCTGCCTACGTGCTCATGATCAAAACGCTCCGTTCCTCCGCC 33177

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db      33176 ATGGACGAGCACTCGGAGAAGAAAAGGCGGAGGGCTATCCGCCTCATCATCACGGTGCTG 33117

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db      33116 TCCATGTACTTCATCTGCTTCGCTCCCAGCAACGTGCTGCTCGTCGTGCATTATTTCCTC 33057

Qy        347 TyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGly 366
Db      33056 ATCAAAAGCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTCGTCGCCCTCTGCCTGTCC 32997

Qy        367 SerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThrArgAsnHis 386
Db      32996 ACCCTCAACAGCTGCATAGACCCCTTTGTCTACTACTTTGTTTCGAAA------------ 32949

Qy        387 SerThrAlaTyrLeuThrLys******AsnAspLeuArgGluGlnGlyGln 403
Db      32948 ------------------------------GATTTCAGGGACCAGGCCAGA 32928

RESULT 27
AC130634/c
LOCUS       AC130634               151166 bp    DNA     linear   HTG 13-AUG-2002
DEFINITION  Rattus norvegicus clone CH230-477A14, *** SEQUENCING IN PROGRESS
            ***, 37 unordered pieces.
ACCESSION   AC130634
VERSION     AC130634.1  GI:22212993
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 151166)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., Mcleod,M., Mcneill,T., Meenen,E., Milosavljevic,A.,
            Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
            Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
            Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G.,
            Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,
            Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A.,
            Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E.,
            Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
            Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,
            Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S.,
            Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
            Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,
            Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,
            Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
            Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D.,
            Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
            Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R.,
```

```
            Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
            Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
            Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
            Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 151166)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-AUG-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
              Center project name: GXIX
              Center clone name: CH230-477A14
            -------------- Summary Statistics
              Sequencing vector: Plasmid;
              Chemistry: Dye-terminator Big Dye: 100% of reads
              Assembly program: Phrap; version 0.990329
              Consensus quality: 115058 bases at least Q40
              Consensus quality: 119558 bases at least Q30
              Consensus quality: 122439 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 37 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1149: contig of 1149 bp in length
            *     1150     1249: gap of unknown length
            *     1250     2920: contig of 1671 bp in length
            *     2921     3020: gap of unknown length
            *     3021     4429: contig of 1409 bp in length
            *     4430     4529: gap of unknown length
            *     4530     5676: contig of 1147 bp in length
            *     5677     5776: gap of unknown length
            *     5777     7247: contig of 1471 bp in length
            *     7248     7347: gap of unknown length
            *     7348     8743: contig of 1396 bp in length
            *     8744     8843: gap of unknown length
            *     8844    10237: contig of 1394 bp in length
            *    10238    10337: gap of unknown length
            *    10338    11514: contig of 1177 bp in length
            *    11515    11614: gap of unknown length
            *    11615    12982: contig of 1368 bp in length
            *    12983    13082: gap of unknown length
            *    13083    14431: contig of 1349 bp in length
            *    14432    14531: gap of unknown length
            *    14532    17021: contig of 2490 bp in length
            *    17022    17121: gap of unknown length
            *    17122    19034: contig of 1913 bp in length
            *    19035    19134: gap of unknown length
            *    19135    20417: contig of 1283 bp in length
            *    20418    20517: gap of unknown length
            *    20518    22781: contig of 2264 bp in length
            *    22782    22881: gap of unknown length
            *    22882    25072: contig of 2191 bp in length
            *    25073    25172: gap of unknown length
            *    25173    26957: contig of 1785 bp in length
            *    26958    27057: gap of unknown length
            *    27058    29399: contig of 2342 bp in length
            *    29400    29499: gap of unknown length
            *    29500    33463: contig of 3964 bp in length
            *    33464    33563: gap of unknown length
            *    33564    35346: contig of 1783 bp in length
            *    35347    35446: gap of unknown length
            *    35447    38703: contig of 3257 bp in length
            *    38704    38803: gap of unknown length
            *    38804    42453: contig of 3650 bp in length
            *    42454    42553: gap of unknown length
            *    42554    45562: contig of 3009 bp in length
            *    45563    45662: gap of unknown length
            *    45663    49065: contig of 3403 bp in length
            *    49066    49165: gap of unknown length
            *    49166    51875: contig of 2710 bp in length
            *    51876    51975: gap of unknown length
            *    51976    54628: contig of 2653 bp in length
            *    54629    54728: gap of unknown length
            *    54729    58389: contig of 3661 bp in length
            *    58390    58489: gap of unknown length
            *    58490    62913: contig of 4424 bp in length
            *    62914    63013: gap of unknown length
            *    63014    68649: contig of 5636 bp in length
            *    68650    68749: gap of unknown length
            *    68750    76294: contig of 7545 bp in length
            *    76295    76394: gap of unknown length
            *    76395    81146: contig of 4752 bp in length
            *    81147    81246: gap of unknown length
            *    81247    86683: contig of 5437 bp in length
            *    86684    86783: gap of unknown length
            *    86784    91620: contig of 4837 bp in length
            *    91621    91720: gap of unknown length
            *    91721    99377: contig of 7657 bp in length
            *    99378    99477: gap of unknown length
            *    99478   108730: contig of 9253 bp in length
            *   108731   108830: gap of unknown length
            *   108831   120614: contig of 11784 bp in length
            *   120615   120714: gap of unknown length
            *   120715   130786: contig of 10072 bp in length
            *   130787   130886: gap of unknown length
            *   130887   151166: contig of 20280 bp in length.
FEATURES             Location/Qualifiers
     source          1. .151166
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-477A14"
BASE COUNT    37400 a  32808 c  33338 g  36965 t  10655 others
ORIGIN

Alignment Scores:
Pred. No.:              5.61e-37        Length:         151166
Score:                  592.50          Matches:        121
Percent Similarity:     54.71%          Conservative:   65
Best Local Similarity:  35.59%          Mismatches:     103
Query Match:            27.47%          Indels:         51
DB:                     2               Gaps:           8

US-09-208-629F-6 (1-408) x AC130634 (1-151166)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              ||||||   :::                  |||   |||||||||
Db      95329 GGGGCTCCAGTTGAACCAGGCTTTTCCGTTGATGAATTCTCTGCATCC---------- 95282

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                              ||||||      |||:::|||      :::|||   |||
Db      95281 -----------------GTCCTCACCGGGAAGCTGACCACCGTCTTTCTCCCGGTCATC 95240

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||::::::|||||||||:::|||:::|||:::|||   :::   |||||:::   |||
Db      95239 TACATCATTGTCTTTGTAATTGGTTTGCCCAGTAATGGTATGGCCCTCTGGGTCTTCTTC 95180

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                 |||   |||      |||||||||:::||||||
Db      95179 TTCCGAACGAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAACCTGGCCTTGGCAGAC 95120

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
```

```
Db      95119 CTCCTCTCTGTCATCTGGTTCCCCCTGAAGATCTCCTACCACCTCCATGGCAACGACTGG 95060

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198

Db      95059 ACCTATGGGGATGCGCTCTGCAAGGTGCTCATTGGCTTTTTCTACGGCAATATGTACTGC 95000

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218

Db      94999 TCCATCCTTTTCATGACCTGCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG 94940

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--------- 235

Db      94939 GGACAC------------TCCAGGAAGAGGGCCAACATCGCTGTTGGCGTCTCCCTGGCC 94892

Qy        236 ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyr 255

Db      94891 ATCTGGCTCCTGATTTTTCTGGTCACCATCCCTCTGTACGTCATGAGGCAGACCATCTAC 94832

Qy        256 LeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSer 275

Db      94831 ATTCCAGCCTTGAACATCACCACCTGTCACGACGTG--------------------CTG 94793

Qy        276 ProPheGlnLeuTyr------------------TyrPheIleSerLeuAlaPhePheGly 289

Db      94792 CCCGAGGAGGTCCTGGTGGGGGACATGTTCAGTTACTTCCTCTCCCTGGCCATTGGAGTC 94733

Qy        290 PheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsn 309

Db      94732 TTTCTGTTCCCAGCCCTCCTTACTGCGTCTGCCTACGTGCTCATGATCAAAACGCTCCGT 94673

Qy        310 Ala------------------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeu 323

Db      94672 TCCTCCGCCATGGACGAGCACTCGGAGAAGAAAAGGCGGAGGGCTATCCGCCTCATCATC 94613

Qy        324 LeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHis 343

Db      94612 ACGGTGCTGTCCATGTACTTCATCTGCTTCGCTCCCAGCAACGTGCTGCTCGTCGTGCAT 94553

Qy        344 HisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeu 363

Db      94552 TATTTCCTCATCAAAAGCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTCGTCGCCCTC 94493

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383

Db      94492 TGCCTGTCCACCCTCAACAGCTGCATAGACCCCTTTGTCTACTACTTTGTTTCGAAA--- 94436

Qy        384 ArgAsnHisSerThrAlaTyrLeuThrLys******AsnAspLeuArgGluGlnGlyGln 403

Db      94435 ------------------------------------GATTTCAGGGACCAGGCCAGA 94415
```

RESULT 28
BC018130
LOCUS BC018130 2876 bp mRNA linear PRI 06-DEC-2001
DEFINITION Homo sapiens, coagulation factor II (thrombin) receptor-like 1, clone MGC:9298 IMAGE:3895653, mRNA, complete cds.
ACCESSION BC018130
VERSION BC018130.1 GI:17390291
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2876)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 14 Row: h Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8051581.

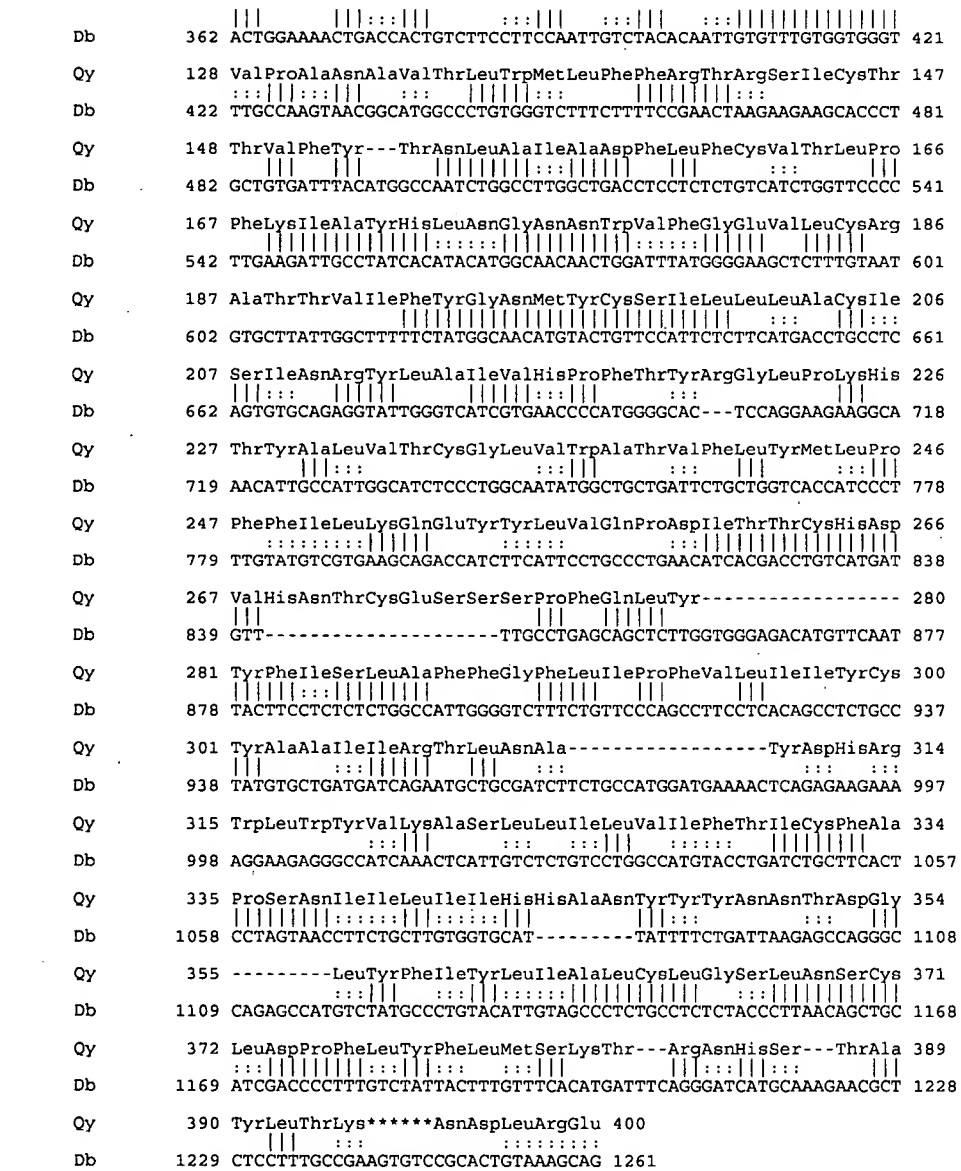
FEATURES Location/Qualifiers
source 1. .2876
/organism="Homo sapiens"
/db_xref="LocusID:2150"
/db_xref="taxon:9606"
/clone="MGC:9298 IMAGE:3895653"
/tissue_type="Pancreas, epithelioid carcinoma"
/clone_lib="NIH_MGC_70"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
CDS 155. .1348
/codon_start=1
/product="coagulation factor II (thrombin) receptor-like 1"
/protein_id="AAH18130.1"
/db_xref="GI:17390292"
/translation="MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG
TSHVTGKGVTVETVFSVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLF
RTKKKHPAVIYMANLALADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMY
CSILFMTCLSVQRYWVIVNPMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFI
PALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDE
NSEKKRKRAIKLIVSVLAMYLICFTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLST
LNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTT
VKTSY"
BASE COUNT 771 a 632 c 629 g 844 t
ORIGIN

Alignment Scores:

Pred. No.:	7.13e-39	Length:	2876
Score:	591.50	Matches:	142
Percent Similarity:	50.85%	Conservative:	67
Best Local Similarity:	34.55%	Mismatches:	145
Query Match:	27.42%	Indels:	57
DB:	9	Gaps:	14

US-09-208-629F-6 (1-408) x BC018130 (1-2876)

```
Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27

Db        146 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 205

Qy         28 AlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47

Db        206 GCCTCTCTC------------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 244

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67

Db        245 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA-------------- 289

Qy         68 GluGluProProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87

Db        290 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 331

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107

Db        332 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 361

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
```



```
RESULT 29
HSU34038
LOCUS       HSU34038                1451 bp    mRNA    linear   PRI 21-FEB-1997
DEFINITION  Human proteinase-activated receptor-2 mRNA, complete cds.
ACCESSION   U34038
VERSION     U34038.1  GI:1041728
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1451)
  AUTHORS   Bohm,S.K., Kong,W., Bromme,D., Smeekens,S.P., Anderson,D.C.,
            Connolly,A., Kahn,M., Nelken,N.A., Coughlin,S.R., Payan,D.G. and
            Bunnett,N.W.
  TITLE     Molecular cloning, expression and potential functions of the human
            proteinase-activated receptor-2
  JOURNAL   Biochem. J. 314 (Pt 3), 1009-1016 (1996)
  MEDLINE   96177879
   PUBMED   8615752
REFERENCE   2  (bases 1 to 1451)
  AUTHORS   Bohm,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-1995) Stephan Bohm, Department of Surgery, School
            of Medicine, University of California at San Francisco, 521
            Parnassus Avenue, San Francisco, CA 94143-0660, USA
FEATURES             Location/Qualifiers
     source          1. .1451
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="kidney"
     5'UTR           1. .147
     CDS             148. .1341
                     /note="G-protein-coupled receptor"
                     /codon_start=1
                     /product="proteinase-activated receptor-2"
                     /protein_id="AAB47871.1"
                     /db_xref="GI:1041729"
                     /translation="MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG
                     TSHVTGKGVTVETVFSVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLF
                     RTKKKHPAVIYMANLALADLLSVIWFPLKIAYHIHANNWIYGEALCNVLIGFFYGNMY
                     CSILFMTCLSVQRYWVIVNPMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFI
                     PALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDE
                     NSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLST
                     LNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTT
                     VKTSY"
     3'UTR           1342. .1451
BASE COUNT      310 a    389 c    346 g    406 t
ORIGIN

Alignment Scores:
Pred. No.:              8.29e-39        Length:         1451
Score:                  586.50          Matches:        141
Percent Similarity:     50.61%          Conservative:   67
Best Local Similarity:  34.31%          Mismatches:     146
Query Match:            27.19%          Indels:         57
DB:                     9               Gaps:           14

US-09-208-629F-6 (1-408) x HSU34038 (1-1451)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
              |||   |||:::|||      :::          :::   |||   |||      |||
Db        139 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 198

Qy         28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
              |||   |||                  |||   ||||||            ||||||
Db        199 GCCTCTCTC-----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 237

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
                 :::|||            |||           |||
Db        238 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA--------------- 282

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
                           |||   :::   |||     |||   |||:::
Db        283 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 324

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
                     |||   |||||||||                                  |||
Db        325 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 354

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
              |||      |||:::|||      :::|||   :::|||   :::||||||||||||||||
Db        355 ACTGGAAAACTGACCACGGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGTGGTGGGT 414
```

```
Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
Db        415 TTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 474

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db        475 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 534

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db        535 TTGAAGATTGCCTATCACATACATGCCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 594

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
Db        595 GTGCTTATTGGCTTTTTCTATGGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 654

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
Db        655 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 711

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
Db        712 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCACCATCCCT 771

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db        772 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 831

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr----------------- 280
Db        832 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 870

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db        871 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 930

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla-----------------TyrAspHisArg 314
Db        931 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 990

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
Db        991 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGCCATGTACCTGATCTGCTTCACT 1050

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db       1051 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1101

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db       1102 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1161

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
Db       1162 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1221

Qy        390 TyrLeuThrLys******AsnAspLeuArgGlu 400
Db       1222 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1254

RESULT 30
HSU36753
LOCUS       HSU36753                1124 bp    DNA     linear   PRI 28-FEB-1996
DEFINITION  Human protease-activated receptor 2 gene, exon 2 and partial cds.
ACCESSION   U36753
VERSION     U36753.1  GI:1208539
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1124)
  AUTHORS   Kahn,M.L. and Coughlin,S.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-1995) Mark L. Kahn, CVRI/Medicine, UCSF, 3rd and
            Parnassus, San Francisco, CA 94143-0524, USA
FEATURES             Location/Qualifiers
     source          1. .1124
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q13"
                     /note="Human genomic DNA in P1 bacteriophage library"
     CDS             <3. .1112
                     /codon_start=1
                     /product="protease-activated receptor 2"
                     /protein_id="AAA90957.1"
                     /db_xref="GI:1208540"
                     /translation="TNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSASVLTGK
                     LTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHPAVIYMANLALADLLSVIWFPL
                     KIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKA
                     NIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLVGDMFNYFLSL
                     AIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKKRKRAIKLIVTVLAMYLICFTPSN
                     LLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYYFVSHDFRDHAKNALLCR
                     SVRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY"
     exon            3. .1110
                     /product="protease-activated receptor 2"
                     /number=2
BASE COUNT      255 a    292 c    238 g    339 t
ORIGIN

Alignment Scores:
Pred. No.:              7.45e-39        Length:         1124
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     9               Gaps:           10

US-09-208-629F-6 (1-408) x HSU36753 (1-1124)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db         72 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 119

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
Db        120 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 161

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db        162 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 221

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db        222 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCGTTGGCGGAC 281

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db        282 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 341

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db        342 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGTAACATGTACTGT 401

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db        402 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 461

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db        462 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 518

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
```

```
Db        519 CTGATTCTTCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 578

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
              :::|||||||||||||||||||||                                |||   |||
Db        579 CTGAACATCACGACCTGCCATGATGTT--------------------TTGCCTGAGCAG 617

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              |||                 ||||||:::|||||||||         ||||||
Db        618 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 677

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
              |||      |||          |||      :::||||||   |||   :::
Db        678 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 737

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                          :::   :::            :::|||     :::   :::|||
Db        738 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 797

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
              ::::::   |||||||||   |||||||||::::::|||::::::|||
Db        798 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 848

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
              |||:::         :::   |||         :::|||   :::|||::::::||||||
Db        849 TATTTTCTGATTAAAAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 908

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   :::|||||||||||:::|||||||||:::|||:::   :::|||
Db        909 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 968

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                 |||:::|||:::      |||   |||   :::           :::::::::
Db        969 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1025

RESULT 31
AR012638
LOCUS       AR012638                1255 bp.    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 3 from patent US 5763575.
ACCESSION   AR012638
VERSION     AR012638.1  GI:3970628
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1255)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Agonist and antagonist peptides of the C140 receptor
  JOURNAL   Patent: US 5763575-A 3 09-JUN-1998;
FEATURES             Location/Qualifiers
     source          1. .1255
                     /organism="unknown"
BASE COUNT      294 a    320 c    260 g    381 t
ORIGIN

Alignment Scores:
Pred. No.:              8.46e-39        Length:         1255
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     6               Gaps:           10

US-09-208-629F-6 (1-408) x AR012638 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||   |||:::                        |||   |||||||||
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                                  ||||||      |||:::|||      :::|||   :::
Db        260 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   :::|||||||||||||||:::|||:::|||   :::   ||||||:::
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                |||   |||      |||||||||:::||||||
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
              |||       :::    |||   ||||||||||||||||:::::||||||||||||
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              ::::::||||||   ||||||                ||||||||||||||||||||
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
              :::                |||          |||:::                :::|||
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
              :::   |||      :::|||   :::::::::||||||      ::::::
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
              :::|||||||||||||||||||||                                |||   |||
Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              |||                 ||||||:::|||||||||         ||||||
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
              |||      |||          |||      :::||||||   |||   :::
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                          :::   :::            :::|||     :::   :::|||
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
              ::::::   |||||||||   |||||||||::::::|||::::::|||
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
              |||:::         :::   |||         :::|||   :::|||::::::||||||
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   :::|||||||||||:::|||||||||:::|||:::   :::|||
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                 |||:::|||:::      |||   |||   :::           :::::::::
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 32
AR171258
LOCUS       AR171258                1255 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 3 from patent US 6297026.
ACCESSION   AR171258
VERSION     AR171258.1  GI:17910208
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1255)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Nucleic acids encoding the C140 receptor
  JOURNAL   Patent: US 6297026-A 3 02-OCT-2001;
```

```
FEATURES             Location/Qualifiers
     source          1. .1255
                     /organism="unknown"
BASE COUNT      294 a    320 c    260 g    381 t
ORIGIN

Alignment Scores:
Pred. No.:              8.46e-39       Length:         1255
Score:                  585.50         Matches:        124
Percent Similarity:     54.87%         Conservative:   62
Best Local Similarity:  36.58%         Mismatches:     114
Query Match:            27.14%         Indels:         39
DB:                     6              Gaps:           10

US-09-208-629F-6 (1-408) x AR171258 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||   |||:::                        |||   |||||||||
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                                  ||||||      |||:::|||      :::|||   :::
Db        260 ------------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   :::|||||||||||||||:::|||:::|||   :::   ||||||:::
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                 |||   |||      |||||||||:::||||||
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
                 |||      :::      |||   ||||||||||||||:::::||||||||||||
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              ::::::||||||   ||||||                  ||||||||||||||||||||
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
              :::          |||         |||:::                  :::|||
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
              :::   |||      :::|||   :::::::::||||||      ::::::
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
              :::||||||||||||||||||||                        |||    |||
Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              |||                   ||||||:::|||||||||         ||||||
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
              |||      |||         |||      :::||||||   |||   :::
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                        :::   :::              :::|||      :::   :::|||
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
                 ::::::   |||||||||   |||||||||:::::|||:::::::|||
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
```

```
              |||:::          :::   |||         :::|||   :::|||::::::||||||
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   :::||||||||||||:::||||||||:::|||:::   :::|||
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                 |||:::|||:::      |||   |||   :::            :::::::::
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 33
I42455
LOCUS       I42455                  1255 bp    DNA     linear   PAT 07-OCT-1997
DEFINITION  Sequence 3 from patent US 5629174.
ACCESSION   I42455
VERSION     I42455.1  GI:2467950
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1255)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Recombinant C140 receptor
  JOURNAL   Patent: US 5629174-A 3 13-MAY-1997;
FEATURES             Location/Qualifiers
     source          1. .1255
                     /organism="unknown"
BASE COUNT      294 a    320 c    260 g    381 t
ORIGIN

Alignment Scores:
Pred. No.:              8.46e-39       Length:         1255
Score:                  585.50         Matches:        124
Percent Similarity:     54.87%         Conservative:   62
Best Local Similarity:  36.58%         Mismatches:     114
Query Match:            27.14%         Indels:         39
DB:                     6              Gaps:           10

US-09-208-629F-6 (1-408) x I42455 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||   |||:::                        |||   |||||||||
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                                  ||||||      |||:::|||      :::|||   :::
Db        260 ------------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   :::|||||||||||||||:::|||:::|||   :::   ||||||:::
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                 |||   |||      |||||||||:::||||||
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
                 |||      :::      |||   ||||||||||||||:::::||||||||||||
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              ::::::||||||   ||||||                  ||||||||||||||||||||
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
              :::          |||         |||:::                  :::|||
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658
```

```
Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 34
I87849
LOCUS       I87849                 1255 bp    DNA     linear   PAT 10-AUG-1998
DEFINITION  Sequence 3 from patent US 5716789.
ACCESSION   I87849
VERSION     I87849.1  GI:3407789
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1255)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Method to determine ligands, agonist and antagonist of C140
            receptor
  JOURNAL   Patent: US 5716789-A 3 10-FEB-1998;
FEATURES             Location/Qualifiers
     source          1. .1255
                     /organism="unknown"
BASE COUNT      294 a    320 c    260 g    381 t
ORIGIN

Alignment Scores:
Pred. No.:              8.46e-39        Length:         1255
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     6               Gaps:           10

US-09-208-629F-6 (1-408) x I87849 (1-1255)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db        212 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 259

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
```

```
Db        260 ------------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 301

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db        302 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 361

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db        362 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db        422 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 481

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db        482 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 541

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db        542 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 601

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db        602 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 658

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db        659 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 718

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db        719 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 757

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db        758 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
Db        818 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 877

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db        878 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 937

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db        938 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 988

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db        989 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1048

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db       1049 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1108

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
Db       1109 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1165

RESULT 35
HSPAR2B
LOCUS       HSPAR2B                1289 bp    DNA     linear   PRI 21-JUL-1997
DEFINITION  H.sapiens partial gene for proteinase-activated receptor 2 (1289
            BP).
ACCESSION   Z49994
VERSION     Z49994.1  GI:1008086
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1289)
  AUTHORS   Nystedt,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-1995) Sverker Nystedt, Division of Neurobiology,
            The Wallenberg, Laboratory, Lund University, Soelvegatan 33A, Lund,
            S-223 62, Sweden
REFERENCE   2  (bases 1 to 1289)
  AUTHORS   Nystedt,S., Emilsson,K., Larsson,A.K., Strombeck,B. and Sundelin,J.
  TITLE     Molecular cloning and functional expression of the gene encoding
            the human proteinase-activated receptor 2
  JOURNAL   Eur. J. Biochem. 232 (1), 84-89 (1995)
  MEDLINE   96048032
   PUBMED   7556175
FEATURES             Location/Qualifiers
     source          1. .1289
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="q13"
     intron          <1. .96
     exon            97. .>1289
                     /product="proteinase-activated receptor 2"
                     /usedin=Z49993:par2_CDS
     3'UTR           1209. .>1289
BASE COUNT      300 a    320 c    270 g    399 t
ORIGIN

Alignment Scores:
Pred. No.:                8.73e-39       Length:          1289
Score:                    585.50         Matches:         124
Percent Similarity:       54.87%         Conservative:    62
Best Local Similarity:    36.58%         Mismatches:      114
Query Match:              27.14%         Indels:          39
DB:                       9              Gaps:            10

US-09-208-629F-6 (1-408) x HSPAR2B (1-1289)

Qy          80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
               |||   |||:::                  |||   |||||||||
Db         168 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 215

Qy         100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                                 ||||||      |||:::|||      :::|||   :::
Db         216 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 257

Qy         120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
               |||   :::|||||||||||||||:::|||:::|||   :::   |||||:::
Db         258 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 317

Qy         140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
               ||||||||||:::                |||   |||      |||||||||:::||||||
Db         318 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 377

Qy         159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
                  |||      :::      |||   ||||||||||||||:::::||||||||||||
Db         378 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 437

Qy         179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
               ::::::||||||   ||||||                |||||||||||||||||||
Db         438 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 497

Qy         199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
               |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db         498 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 557

Qy         219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
                  :::               |||         |||:::               :::|||
Db         558 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 614

Qy         239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
                  :::   |||      :::|||   :::::::::||||||      ::::::
Db         615 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 674

Qy         259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
                  :::|||||||||||||||||||||                          |||   |||
Db         675 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 713

Qy         279 LeuTyr--------------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
               |||                     ||||||:::|||||||||        ||||||
Db         714 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 773

Qy         293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
               |||      |||              |||      :::||||||   |||   :::
Db         774 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 833

Qy         311 ------------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                         :::   :::          :::|||      :::   :::|||
Db         834 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 893

Qy         327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
                  ::::::   |||||||||   |||||||||::::::|||::::::|||
Db         894 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 944

Qy         347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
               |||:::         :::   |||         :::|||   :::|||::::::||||||
Db         945 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 1004

Qy         364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
               ||||||   :::||||||||||||:::||||||||:::|||:::   :::|||
Db        1005 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 1064

Qy         384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                  |||:::|||:::      |||   |||   :::           :::::::::
Db        1065 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 1121

RESULT 36
AF400075
LOCUS       AF400075               18351 bp    DNA     linear   PRI 18-JAN-2002
DEFINITION  Homo sapiens coagulation factor II (thrombin) receptor-like 1
            (F2RL1) gene, complete cds.
ACCESSION   AF400075
VERSION     AF400075.1  GI:15021772
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 18351)
  AUTHORS   Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
            Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2001) Molecular Biotechnology, University of
            Washington, 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI Program for
            Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.mbt.washington.edu)
            This sequence consists of 2 contigs.  The gap between the contigs
            is represented as a run of N.
                1    10774: contig of 10774 bp in length
            10775    10874: gap of unknown length
            10875    18351: contig of 7477 bp in length.
FEATURES             Location/Qualifiers
     source          1. .18351
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     repeat_region   343. .492
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     variation       467
                     /frequency="0.09"
                     /replace="a"
     variation       531
                     /frequency="0.01"
                     /replace="g"
     variation       678
```

```
                /frequency="0.01"
                /replace="c"
repeat_region   728. .780
                /rpt_family="L2"
                /rpt_type=dispersed
variation       829
                /frequency="0.03"
                /replace="a"
repeat_region   913. .1028
                /rpt_family="FLAM_C"
                /rpt_type=dispersed
variation       925
                /frequency="0.01"
                /replace="t"
repeat_region   1059. .1214
                /rpt_family="L3"
                /rpt_type=dispersed
repeat_region   1292. .1578
                /rpt_family="AluSg"
                /rpt_type=dispersed
variation       1538
                /frequency="0.01"
                /replace="g"
variation       1622
                /frequency="0.02"
                /replace="t"
variation       1623
                /frequency="0.01"
                /replace="g"
repeat_region   1795. .1926
                /rpt_family="MIR"
                /rpt_type=dispersed
variation       2122
                /frequency="0.01"
                /replace="g"
variation       2254
                /frequency="0.01"
                /replace="c"
variation       2457
                /frequency="0.10"
                /replace="g"
gene            2462. .17176
                /gene="F2RL1"
mRNA            join(2462. .2687,15958. .17176)
                /gene="F2RL1"
                /product="coagulation factor II (thrombin) receptor-like
                1"
variation       2561
                /gene="F2RL1"
                /frequency="0.28"
                /replace="c"
CDS             join(2606. .2687,15958. .17069)
                /gene="F2RL1"
                /codon_start=1
                /product="coagulation factor II (thrombin) receptor-like
                1"
                /protein_id="AAK77914.1"
                /db_xref="GI:15021773"
                /translation="MRSPSAAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG
                TSHVTGKGVTVETVFSVDEFSASVLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLF
                RTKKKHPAVIYMANLALADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMY
                CSILFMTCLSVQRYWVIVNPMGHSRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFI
                PALNITTCHDVLPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDE
                NSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVVHYFLIKSQGQSHVYALYIVALCLST
                LNSCIDPFVYYFVSHDFRDHAKNALLCRSVRTVKQMQVSLTSKKHSRKSSSYSSSSTT
                VKTSY"
variation       2667
                /gene="F2RL1"
                /frequency="0.02"
                /replace="t"
variation       2770
                /gene="F2RL1"
                /frequency="0.09"
                /replace="a"
variation       3069
                /gene="F2RL1"
                /frequency="0.03"
                /replace="a"
variation       3119
                /gene="F2RL1"
                /frequency="0.32"
                /replace="t"
variation       3151
                /gene="F2RL1"
                /frequency="0.01"
                /replace="g"
variation       3263
                /gene="F2RL1"
                /frequency="0.33"
                /replace="g"
variation       3306
                /gene="F2RL1"
                /frequency="0.02"
                /replace="t"
variation       3330
                /gene="F2RL1"
                /frequency="0.01"
                /replace="c"
variation       3331
                /gene="F2RL1"
                /frequency="0.07"
                /replace="t"
variation       3353
                /gene="F2RL1"
                /frequency="0.01"
                /replace="a"
variation       3417
                /gene="F2RL1"
                /frequency="0.01"
                /replace="c"
variation       4160
                /gene="F2RL1"
                /frequency="0.01"
                /replace="c"
variation       4456
                /gene="F2RL1"
                /frequency="0.01"
                /replace=""
variation       4500
                /gene="F2RL1"
                /frequency="0.02"
                /replace="t"
variation       4663
                /gene="F2RL1"
                /frequency="0.46"
                /replace="t"
repeat_region   4710. .4817
                /rpt_family="L2"
                /rpt_type=dispersed
variation       4715
                /gene="F2RL1"
                /frequency="0.15"
                /replace="g"
variation       4972
                /gene="F2RL1"
                /frequency="0.02"
                /replace="c"
repeat_region   4993. .5299
                /rpt_family="AluSx"
                /rpt_type=dispersed
variation       5065
                /gene="F2RL1"
                /frequency="0.01"
                /replace="g"
variation       5159
                /gene="F2RL1"
```

```
                    /frequency="0.01"
                    /replace=""
     variation      5295
                    /gene="F2RL1"
                    /frequency="0.01"
                    /replace="t"
     variation      5296
                    /gene="F2RL1"
                    /frequency="0.01"
                    /replace="a"
     variation      5366
                    /gene="F2RL1"
                    /frequency="0.17"
                    /replace="t"
     repeat_region  5643. .5940
                    /rpt_family="AluSx"

Alignment Scores:
Pred. No.:              1.85e-37        Length:         18351
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     9               Gaps:           10

US-09-208-629F-6 (1-408) x AF400075 (1-18351)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db      16029 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 16076

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
Db      16077 ------------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 16118

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db      16119 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 16178

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db      16179 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 16238

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db      16239 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 16298

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db      16299 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 16358

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db      16359 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 16418

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db      16419 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 16475

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db      16476 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 16535

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db      16536 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 16574

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db      16575 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 16634

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
Db      16635 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 16694

Qy        311 ------------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db      16695 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 16754

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db      16755 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 16805

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db      16806 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 16865

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db      16866 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 16925

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
Db      16926 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 16982

RESULT 37
AC010621
LOCUS       AC010621               52368 bp    DNA     linear   PRI 26-JAN-2002
DEFINITION  Homo sapiens chromosome 5 clone CTD-2034J20, complete sequence.
ACCESSION   AC010621
VERSION     AC010621.7  GI:18376859
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52368)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 52368)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 52368)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jan 26, 2002 this sequence version replaced gi:13699549.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.4% of Sequence;
            Estimated Total Number of Errors is 0.2.
            NOTE:  This insert is not the entire sequence of the clone (entire
            sequence is 142.2kb). It is clipped at the overlap with AC025188.
            The number of bases overlapped is 12615.
FEATURES             Location/Qualifiers
     source          1. .52368
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2034J20"
BASE COUNT    13910 a  11840 c  12216 g  14402 t
ORIGIN

Alignment Scores:
Pred. No.:              6.16e-37        Length:         52368
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     9               Gaps:           10
```

US-09-208-629F-6 (1-408) x AC010621 (1-52368)

```
Qy        80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
             |||   |||:::                    |||   |||||||||
Db     43609 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 43656

Qy       100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                             ||||||      |||:::|||      :::|||   :::
Db     43657 ------------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 43698

Qy       120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
             |||   :::|||||||||||||||||:::|||:::|||   :::   ||||||:::
Db     43699 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 43758

Qy       140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
             |||||||||:::                 |||   |||      |||||||||:::||||||
Db     43759 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 43818

Qy       159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
             |||      :::      |||   ||||||||||||||||:::::|||||||||||||
Db     43819 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 43878

Qy       179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
             ::::::||||||   ||||||                  ||||||||||||||||||||
Db     43879 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 43938

Qy       199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
             |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db     43939 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 43998

Qy       219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
                :::             |||         |||:::                :::|||
Db     43999 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 44055

Qy       239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
                :::   |||      :::|||   :::::::::||||||      ::::::
Db     44056 CTGATTCTGCTGGTCACCATTCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 44115

Qy       259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
              :::||||||||||||||||||||                          |||   |||
Db     44116 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 44154

Qy       279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
             |||                    ||||||:::|||||||||         ||||||
Db     44155 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 44214

Qy       293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
             |||      |||         |||      :::||||||   |||   :::
Db     44215 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 44274

Qy       311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                        :::   :::                :::|||      :::   :::|||
Db     44275 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 44334

Qy       327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
                :::::   |||||||||   |||||||||:::::|||:::::|||
Db     44335 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 44385

Qy       347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
             |||:::      :::   |||         :::|||   :::|||::::::||||||
Db     44386 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 44445

Qy       364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
             ||||||   :::|||||||||||:::||||||||||:::|||:::   :::|||
Db     44446 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 44505

Qy       384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                |||:::|||:::      |||   |||   :::           :::::::::
Db     44506 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 44562

RESULT 38
AC068682/c
```

```
LOCUS       AC068682              184536 bp    DNA     linear   HTG 23-SEP-2000
DEFINITION  Homo sapiens clone RP11-206N2, WORKING DRAFT SEQUENCE, 37 unordered
            pieces.
ACCESSION   AC068682
VERSION     AC068682.3  GI:10280868
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 184536)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-206N2
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 184536)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 23, 2000 this sequence version replaced gi:8247861.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L5802
                Center clone name: 206_N_2
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 167659 bases at least Q40
                Consensus quality: 175736 bases at least Q30
                Consensus quality: 178914 bases at least Q20
                Insert size: 154000; agarose-fp
                Insert size: 180936; sum-of-contigs
                Quality coverage: 5.0 in Q20 bases; agarose-fp
                Quality coverage: 4.3 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 37 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1154: contig of 1154 bp in length
            *    1155 1254: gap of      100 bp
```

```
*     1255     2766: contig of 1512 bp in length
*     2767 2866: gap of      100 bp
*     2867     4073: contig of 1207 bp in length
*     4074 4173: gap of      100 bp
*     4174     5191: contig of 1018 bp in length
*     5192 5291: gap of      100 bp
*     5292     7387: contig of 2096 bp in length
*     7388 7487: gap of      100 bp
*     7488     9087: contig of 1600 bp in length
*     9088 9187: gap of      100 bp
*     9188    10361: contig of 1174 bp in length
*    10362 10461: gap of      100 bp
*    10462    12065: contig of 1604 bp in length
*    12066 12165: gap of      100 bp
*    12166    13498: contig of 1333 bp in length
*    13499 13598: gap of      100 bp
*    13599    15085: contig of 1487 bp in length
*    15086 15185: gap of      100 bp
*    15186    17249: contig of 2064 bp in length
*    17250 17349: gap of      100 bp
*    17350    19243: contig of 1894 bp in length
*    19244 19343: gap of      100 bp
*    19344    21127: contig of 1784 bp in length
*    21128 21227: gap of      100 bp
*    21228    23345: contig of 2118 bp in length
*    23346 23445: gap of      100 bp
*    23446    25965: contig of 2520 bp in length
*    25966 26065: gap of      100 bp
*    26066    28195: contig of 2130 bp in length
*    28196 28295: gap of      100 bp
*    28296    32216: contig of 3921 bp in length
*    32217 32316: gap of      100 bp
*    32317    36954: contig of 4638 bp in length
*    36955 37054: gap of      100 bp
*    37055    40348: contig of 3294 bp in length
*    40349 40448: gap of      100 bp
*    40449    43556: contig of 3108 bp in length
*    43557 43656: gap of      100 bp
*    43657    47926: contig of 4270 bp in length
*    47927 48026: gap of      100 bp
*    48027    54079: contig of 6053 bp in length
*    54080 54179: gap of      100 bp
*    54180    58962: contig of 4783 bp in length
*    58963 59062: gap of      100 bp
*    59063    63739: contig of 4677 bp in length
*    63740 63839: gap of      100 bp
*    63840    71551: contig of 7712 bp in length
*    71552 71651: gap of      100 bp
*    71652    78218: contig of 6567 bp in length
*    78219 78318: gap of      100 bp
*    78319    84982: contig of 6664 bp in length
*    84983 85082: gap of      100 bp
*    85083    91930: contig of 6848 bp in length
*    91931 92030: gap of      100 bp
*    92031    98922: contig of 6892 bp in length
*    98923 99022: gap of      100 bp
*    99023   106496: contig of 7474 bp in length
*   106497 106596: gap of      100 bp
*   106597   114950: contig of 8354 bp in length
*   114951 115050: gap of      100 bp
*   115051   124353: contig of 9303 bp in length
*   124354 124453: gap of      100 bp
*   124454   132639: contig of 8186 bp in length
*   132640 132739: gap of      100 bp
*   132740   141572: contig of 8833 bp in length
*   141573 141672: gap of      100 bp
*   141673   151302: contig of 9630 bp in length
*   151303 151402: gap of      100 bp
*   151403   161767: contig of 10365 bp in length
*   161768 161867: gap of      100 bp
*   161868   184536: contig of 22669 bp in length.
FEATURES             Location/Qualifiers
     source          1. .184536
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-206N2"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1154
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    1255. .2766
                     /note="assembly_fragment"
     misc_feature    2867. .4073
                     /note="assembly_fragment"
     misc_feature    4174. .5191
                     /note="assembly_fragment"
     misc_feature    5292. .7387
                     /note="assembly_fragment"
     misc_feature    7488. .9087
                     /note="assembly_fragment"
     misc_feature    9188. .10361
                     /note="assembly_fragment"
     misc_feature    10462. .12065
                     /note="assembly_fragment"
     misc_feature    12166. .13498
                     /note="assembly_fragment"
     misc_feature    13599. .15085
                     /note="assembly_fragment"
     misc_feature    15186. .17249
                     /note="assembly_fragment"
     misc_feature    17350. .19243
                     /note="assembly_fragment"
     misc_feature    19344. .21127
                     /note="assembly_fragment"
     misc_feature    21228. .23345
                     /note="assembly_fragment"
     misc_feature    23446. .25965
                     /note="assembly_fragment"
     misc_feature    26066. .28195
                     /note="assembly_fragment"
     misc_feature    28296. .32216
                     /note="assembly_fragment"
     misc_feature    32317. .36954
                     /note="assembly_fragment"
     misc_feature    37055. .40348
                     /note="assembly_fragment"
     misc_feature    40449. .43556
                     /note="assembly_fragment"
     misc_feature    43657. .47926
                     /note="assembly_fragment"
     misc_feature    48027. .54079
                     /note="assembly_fragment"
     misc_feature    54180. .58962
                     /note="assembly_fragment"
     misc_feature    59063. .63739
                     /note="assembly_fragment"

Alignment Scores:
Pred. No.:              2.62e-36        Length:         184536
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     2               Gaps:           10

US-09-208-629F-6 (1-408) x AC068682 (1-184536)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||   |||:::                 |||   |||||||||
Db      66691 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 66644

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
                                  ||||||      |||:::|||      :::|||   :::
Db      66643 -----------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 66602
```

```
Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
              |||   :::|||||||||||||||||:::|||:::|||   :::   ||||||:::
Db      66601 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 66542

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
              |||||||||:::                 |||   |||      |||||||||:::||||||
Db      66541 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 66482

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
              |||      :::      |||   ||||||||||||||||:::::|||||||||||
Db      66481 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 66422

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
              ::::::||||||   ||||||                ||||||||||||||||||||
Db      66421 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 66362

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
              |||||||||   :::   |||:::|||:::   ||||||      ||||||:::|||
Db      66361 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 66302

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
              :::               |||         |||:::                 :::|||
Db      66301 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 66245

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
              :::   |||      :::|||   :::::::::||||||      ::::::
Db      66244 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 66185

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
              :::|||||||||||||||||||||                          |||   |||
Db      66184 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 66146

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
              |||                   ||||||:::|||||||||         ||||||
Db      66145 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 66086

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
              |||      |||         |||      :::||||||   |||   :::
Db      66085 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 66026

Qy        311 -----------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
                        :::   :::            :::|||      :::   :::|||
Db      66025 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 65966

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
              ::::::   |||||||||   |||||||||::::::|||:::::|||
Db      65965 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 65915

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
              |||:::          :::   |||         :::|||   :::|||::::::||||||
Db      65914 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 65855

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
              ||||||   :::|||||||||||||:::||||||||||:::|||:::   :::|||
Db      65854 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 65795

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
                 |||:::|||:::      |||   |||   :::              :::::::::
Db      65794 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 65738

RESULT 39
AC114962
LOCUS       AC114962               204485 bp    DNA     linear   HTG 14-MAR-2002
DEFINITION  Homo sapiens chromosome 5 clone RP11-206N2, WORKING DRAFT SEQUENCE,
            8 unordered pieces.
ACCESSION   AC114962
VERSION     AC114962.1  GI:19424423
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 204485)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 204485)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 485783
            Center clone name: RPCI-11_206N2
            -------
            Summary Statistics
            Consensus quality: 195215 bases at least Q40
            Consensus quality: 199004 bases at least Q30
            Consensus quality: 200415 bases at least Q20
            Estimated insert size: 175000; agarose-fp estimation
            Estimated insert size: 203785; sum-of-contigs estimation
            Quality coverage:  9.78 in Q20 bases; agarose-fp estimation
            Quality coverage:  8.4 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1058: contig of 1058 bp in length
            *     1059     1158: gap of unknown length
            *     1159     2342: contig of 1184 bp in length
            *     2343     2442: gap of unknown length
            *     2443     4907: contig of 2465 bp in length
            *     4908     5007: gap of unknown length
            *     5008    19885: contig of 14878 bp in length
            *    19886    19985: gap of unknown length
            *    19986    35850: contig of 15865 bp in length
            *    35851    35950: gap of unknown length
            *    35951    53802: contig of 17852 bp in length
            *    53803    53902: gap of unknown length
            *    53903   107788: contig of 53886 bp in length
            *   107789   107888: gap of unknown length
            *   107889   204485: contig of 96597 bp in length.
FEATURES             Location/Qualifiers
     source          1. .204485
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-206N2"
                     /clone_lib="RPCI human BAC library 11"
BASE COUNT    56752 a  44584 c  44954 g  57495 t    700 others
ORIGIN

Alignment Scores:
Pred. No.:              2.95e-36        Length:         204485
Score:                  585.50          Matches:        124
Percent Similarity:     54.87%          Conservative:   62
Best Local Similarity:  36.58%          Mismatches:     114
Query Match:            27.14%          Indels:         39
DB:                     2               Gaps:           10

US-09-208-629F-6 (1-408) x AC114962 (1-204485)

Qy         80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
              |||   |||:::                     |||    |||||||||
Db      73897 GGAGTTACAGTTGAAACAGTCTTTTCTGTGGATGAGTTTTCTGCATCT------------ 73944

Qy        100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
```

```
Db      73945 ------------------GTCCTCACTGGAAAACTGACCACTGTCTTCCTTCCAATTGTC 73986

Qy        120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db      73987 TACACAATTGTGTTTGTGGTGGGTTTGCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTT 74046

Qy        140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db      74047 TTCCGAACTAAGAAGAAGCACCCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 74106

Qy        159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db      74107 CTCCTCTCTGTCATCTGGTTCCCCTTGAAGATTGCCTATCACATACATGGCAACAACTGG 74166

Qy        179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db      74167 ATTTATGGGGAAGCTCTTTGTAATGTGCTTATTGGCTTTTTCTATGGCAACATGTACTGT 74226

Qy        199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db      74227 TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATG 74286

Qy        219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db      74287 GGGCAC---TCCAGGAAGAAGGCAAACATTGCCATTGGCATCTCCCTGGCAATATGGCTG 74343

Qy        239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db      74344 CTGATTCTGCTGGTCACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCC 74403

Qy        259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db      74404 CTGAACATCACGACCTGTCATGATGTT--------------------TTGCCTGAGCAG 74442

Qy        279 LeuTyr-----------------TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db      74443 CTCTTGGTGGGAGACATGTTCAATTACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 74502

Qy        293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla------ 310
Db      74503 CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTGCGATCTTCTGCC 74562

Qy        311 ------------TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db      74563 ATGGATGAAAACTCAGAGAAGAAAAGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTG 74622

Qy        327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db      74623 GCCATGTACCTGATCTGCTTCACTCCTAGTAACCTTCTGCTTGTGGTGCAT--------- 74673

Qy        347 TyrTyrTyrAsnAsnThrAspGly---------LeuTyrPheIleTyrLeuIleAlaLeu 363
Db      74674 TATTTTCTGATTAAGAGCCAGGGCCAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTC 74733

Qy        364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db      74734 TGCCTCTCTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTCACATGAT 74793

Qy        384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys******AsnAspLeuArgGlu 400
Db      74794 TTCAGGGATCATGCAAAGAACGCTCTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAG 74850

RESULT 40
AR012640
LOCUS       AR012640               1414 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 62 from patent US 5763575.
ACCESSION   AR012640
VERSION     AR012640.1  GI:3970630
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.

REFERENCE   1  (bases 1 to 1414)
  AUTHORS   Sundelin,J. and Scarborough,R.M.
  TITLE     Agonist and antagonist peptides of the C140 receptor
  JOURNAL   Patent: US 5763575-A 62 09-JUN-1998;
FEATURES             Location/Qualifiers
     source          1. .1414
                     /organism="unknown"
BASE COUNT      335 a    361 c    309 g    409 t
ORIGIN

Alignment Scores:
Pred. No.:                1.11e-37      Length:          1414
Score:                    572.50        Matches:         140
Percent Similarity:       49.64%        Conservative:    66
Best Local Similarity:    33.73%        Mismatches:      152
Query Match:              26.54%        Indels:          57
DB:                       6             Gaps:            14

US-09-208-629F-6 (1-408) x AR012640 (1-1414)

Qy          8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
Db         41 TCCAGGAGGATGCGGAGCCCCAGCGCGGCGTGGCTGCTGGGGGCCGCCATCCTGCTAGCA 100

Qy         28 AlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db        101 GCCTCTCTC----------------TCCTGC---AGTGGCACCATCCAAGGAACCAAT 139

Qy         48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
Db        140 AGATCCTCTAAAGGAAGAAGCCTTATTGGTAAGGTTGATGGCACA-------------- 184

Qy         68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
Db        185 --------------TCCCACGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226

Qy         88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db        227 TCTGTGGATGAGTTTTCTGCATCT----------------------------GTCCTC 256

Qy        108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
Db        257 GCTGGAAAACTGACCACTGTCTTCCTTCCAATTGTCTACACAATTGTGTTTGCGGTGGGT 316

Qy        128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
Db        317 TTGCCAAGTAACGGCATGGCCCTATGGGTCTTTCTTTTCCGAACTAAGAAGAAGCACCCT 376

Qy        148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db        377 GCTGTGATTTACATGGCCAATCTGGCCTTGGCTGACCTCCTCTCTGTCATCTGGTTCCCC 436

Qy        167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db        437 TTGAAGATTGCCTATCACATACATGGCAACAACTGGATTTATGGGGAAGCTCTTTGTAAT 496

Qy        187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
Db        497 GTGCTTATTGGCTTTTTCTATCGCAACATGTACTGTTCCATTCTCTTCATGACCTGCCTC 556

Qy        207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
Db        557 AGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCAC---TCCAGGAAGAAGGCA 613

Qy        227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
Db        614 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCT 673

Qy        247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db        674 TTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAACATCACGACCTGTCATGAT 733

Qy        267 ValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr------------------ 280
```

```
                ||||                          |||   ||||||
Db        734 GTT--------------------TTGCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772

Qy        281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
                ||||||:::|||||||||          ||||||   |||      |||
Db        773 TACTTCCTCTCTCTGGCCATTGGGGTCTTTCTGTTCCCAGCCTTCCTCACAGCCTCTGCC 832

Qy        301 TyrAlaAlaIleIleArgThrLeuAsnAla-----------------TyrAspHisArg 314
                |||      :::|||||||   |||   :::                    :::   :::
Db        833 TATGTGCTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892

Qy        315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
                    :::|||      :::   :::|||   ::::::   |||||||||
Db        893 AGGAAGAGGGCCATCAAACTCATTGTCACTGTCCTGGGCATGTACCTGATCTGCTTCACT 952

Qy        335 ProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
                |||||||||:::::::|||:::::::|||          |||:::          :::   |||
Db        953 CCTAGTAACCTTCTGCTTGTGGTGCAT---------TATTTTCTGATTAAGAGCCAGGGC 1003

Qy        355 ---------LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
                      :::|||   :::|||::::::|||||||||||||   :::|||||||||||||
Db       1004 CAGAGCCATGTCTATGCCCTGTACATTGTAGCCCTCTGCCTCTCTACCCTTAACAGCTGC 1063

Qy        372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
                :::||||||||||:::|||:::   :::|||          |||:::|||:::      |||
Db       1064 ATCGACCCCTTTGTCTATTACTTTGTTTCACATGATTTCAGGGATCATGCAAAGAACGCT 1123

Qy        390 TyrLeuThrLys******AsnAspLeuArgGluGlnGlyGlnPro 404
                |||   :::            :::::::::         |||
Db       1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168

Search completed: June 29, 2003, 09:48:53
Job time : 2761.78 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:58:31 ; Search time 51.5632 Seconds
(without alignments)
1630.373 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIKM.....AYLTKXXNDLREQGQPSQRT 408

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	605.5	28.1	399	11	Q8R3I1	Q8r3i1 mus musculu
2	453	21.0	357	13	Q9DE05	Q9de05 raja erinac
3	432.5	20.1	374	13	O57466	O57466 meleagris g
4	422.5	19.6	361	13	Q90X57	Q90x57 xenopus lae
5	387	17.9	359	11	Q99MT7	Q99mt7 mus musculu
6	368.5	17.1	358	4	Q9BY21	Q9by21 homo sapien
7	368.5	17.1	358	4	Q96JZ8	Q96jz8 homo sapien
8	367.5	17.0	309	11	Q8R528	Q8r528 mus musculu
9	367.5	17.0	359	13	Q9PVY7	Q9pvy7 anguilla an
10	358.5	16.6	342	6	Q9GK76	Q9gk76 capra hircu
11	356.5	16.5	342	6	Q9TTY5	Q9tty5 bos taurus
12	353	16.4	367	11	Q9R1V0	Q9r1v0 mus musculu
13	351	16.3	337	4	Q96P68	Q96p68 homo sapien
14	347	16.1	361	6	O46685	O46685 bos taurus
15	345	16.0	359	11	Q9EPP3	Q9epp3 cavia porce
16	341.5	15.8	296	6	Q9TTY6	Q9tty6 canis famil
17	340	15.8	328	11	Q9ERK9	Q9erk9 mus musculu
18	340	15.8	359	6	Q9N0U1	Q9n0u1 ovis aries
19	338	15.7	334	6	O97724	O97724 felis silve
20	337.5	15.6	298	11	Q8VE54	Q8ve54 mus musculu
21	337	15.6	390	13	Q8QGQ4	Q8qgq4 carassius a
22	336.5	15.6	342	6	Q9XSD4	Q9xsd4 sus scrofa
23	334	15.5	359	6	Q9GLN9	Q9gln9 pan troglod
24	333.5	15.5	339	6	Q9TUS6	Q9tus6 papio cynoc
25	333	15.4	353	13	Q8QFR5	Q8qfr5 xenopus lae
26	332	15.4	380	4	Q9UNW8	Q9unw8 homo sapien
27	331.5	15.4	352	6	Q9TV48	Q9tv48 cercopithec
28	331.5	15.4	382	11	Q9Z282	Q9z282 mus musculu
29	331	15.3	359	4	Q8TBK4	Q8tbk4 homo sapien
30	330.5	15.3	339	6	Q9TUR6	Q9tur6 cercopithec
31	330	15.3	342	13	O93239	O93239 cyprinus ca
32	329.5	15.3	339	6	Q9TUU5	Q9tuu5 macaca fusc
33	329.5	15.3	339	6	Q9TUR4	Q9tur4 mandrillus
34	329.5	15.3	352	6	Q95NC3	Q95nc3 miopithecus
35	328.5	15.2	339	6	Q9TUU8	Q9tuu8 cercopithec
36	328.5	15.2	339	6	Q9TSQ3	Q9tsq3 cercopithec
37	328.5	15.2	343	6	Q9N0Z0	Q9n0z0 cercocebus
38	328	15.2	355	6	Q9MYJ8	Q9myj8 callithrix
39	327.5	15.2	339	6	Q9TUU6	Q9tuu6 macaca fusc
40	327.5	15.2	339	6	Q9TQX3	Q9tqx3 mandrillus
41	327.5	15.2	352	6	Q95ND2	Q95nd2 mandrillus
42	327	15.2	337	11	Q921N3	Q921n3 mus musculu
43	326.5	15.1	343	6	Q9BDS6	Q9bds6 macaca fasc
44	325.5	15.1	339	6	Q9TUQ8	Q9tuq8 cercopithec
45	325.5	15.1	342	6	Q9TV16	Q9tv16 pan troglod

ALIGNMENTS

RESULT 1
Q8R3I1
ID Q8R3I1 PRELIMINARY; PRT; 399 AA.
AC Q8R3I1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Coagulation factor II (thrombin) receptor-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC025432; AAH25432.1; -.
KW Receptor.
SQ SEQUENCE 399 AA; 44679 MW; 807C79464AB9B3EF CRC64;

Query Match 28.1%; Score 605.5; DB 11; Length 399;
Best Local Similarity 35.4%; Pred. No. 4.2e-44;
Matches 137; Conservative 68; Mismatches 127; Indels 55; Gaps 12;

```
Qy         22 ALIFAAAGLLLLLPTFCQSGMEN--DTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEGWT 79
              :| :   |: ||  :   |  ||     | :|    |      ||              |
Db          5 SLAWLLGGITLLAASVSCSRTENLAPGRNNSKGRSLIGRLETQPP------------IT 51

Qy         80 GATITVK--IKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWM 137
              |  : |:       | |||          ||  |:|  :| :|::|||:|:|:| : ||:
Db         52 GKGVPVEPGFSIDEFSAS----------ILTGKLTTVFLPVVYIIVFVIGLPSNGMALWI 101

Qy        138 LFFRTRSICTTVFY-TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNM 196
                 |||:     | |  |||:|| |  :  |  |||||:|||||:|| ||:     |||||
Db        102 FLFRTKKKHPAVIYMANLALADLLSVIWFPLAIAYHLHGNNWVYGEALCKVLIGFFYGNM 161

Qy        197 YCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKQE 253
              ||||| : |:|: ||  ||:|    | |:   | :  |:   :|  :||  :| :::||
Db        162 YCSILFMTCLSVQRYWVIVNPM---GHPRKK-ANIAVGVSLAIWLLIFLVTIPLYVMKQT 217
```

```
Qy        254 YYLVQPDITTCHDVHNTCESSSPFQLY------YFISLAFFGFLIPFVLIIYCYAAIIRT 307
Db        218 IYIPALNITTCHDV-------LPEEVLVGDMFNYFLSLAIGVFLFPAILTASAYVLMIKT 270

Qy        308 LNA------YDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLI 361
Db        271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNLLLVVHYFLIKTQRQSHVYALYLV 330

Qy        362 ALCLGSLNSCLDPFLYFLMSKT-RNHS 387
Db        331 ALCLSTLNSCIDPFVYYFVSKDFRDHA 357
```

RESULT 2
Q9DE05
ID Q9DE05 PRELIMINARY; PRT; 357 AA.
AC Q9DE05;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE P2Y receptor.
OS Raja erinacea (Little skate).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=7782;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20459151; PubMed=10900200;
RA Dranoff J.A., O'Neill A.F., Franco A.M., Cai S.Y., Connolly G.C.,
RA Ballatori N., Boyer J.L., Nathanson M.H.;
RT "A primitive ATP receptor from the little skate Raja erinacea.";
RL J. Biol. Chem. 275:30701-30706(2000).
DR EMBL; AF242850; AAG42684.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 357 AA; 41239 MW; 14604EE15DCBDB41 CRC64;

Query Match 21.0%; Score 453; DB 13; Length 357;
Best Local Similarity 32.9%; Pred. No. 4.8e-31;
Matches 102; Conservative 54; Mismatches 132; Indels 22; Gaps 8;

```
Qy        116 IPAIYLLVFVVGVPANAVTLWMLFFRTRSICT-TVFYTNLAIADFLFCVTLPFKIAYHLN 174
Db         38 LPIMYIIVFVTGFIGNSVALWMFIFHMRPWSSITIYMFNLVLADLFYVFSLPILIFYYFN 97

Qy        175 GNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCG 234
Db         98 KTDWIFGELLCKLXRFIFHVNLYGSILFLTCISVHRYTGVVHPMKSLGRLKKKSATIVCV 157

Qy        235 LVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPF 294
Db        158 CVWITVMAGISPILYFSRT-GLRRNKTNTCYD--TTSKELLETYFIYSMSTTFFGFCIPF 214

Qy        295 VLIIYCYAAIIRTLNAYDHRWLW---YVKASLLILVIFTICFAPSNIILIIH-HANYYYN 350
Db        215 ATILVCYGFIVKALISNDMKTPLRGKSVRLVIIVLAVFAISYLPFHVMKNLNLQSRLYYQ 274

Qy        351 NTD------GLYFIYLIALCLGSLNSCLDPFLYFLMSKT-----RNHSTAYLTKXXN--- 396
Db        275 GLDTCEWNRRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRRFTNAASRFMTRRSELSM 334

Qy        397 DLREQGQPSQ 406
Db        335 QFRSEDSPLQ 344

RESULT 3
O57466
ID O57466 PRELIMINARY; PRT; 374 AA.
AC O57466;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE G protein coupled P2Y nucleotide receptor.
GN TP2Y.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=98086419; PubMed=9415702;
RA Boyer J.L., Waldo G.L., Harden T.K.;
RT "Molecular cloning and expression of an avian G protein-coupled P2Y
RT receptor.";
RL Mol. Pharmacol. 52:928-934(1997).
DR EMBL; AF031897; AAC60339.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 374 AA; 42594 MW; 849C465722BDD02B CRC64;

Query Match 20.1%; Score 432.5; DB 13; Length 374;
Best Local Similarity 30.0%; Pred. No. 2.9e-29;
Matches 106; Conservative 57; Mismatches 123; Indels 67; Gaps 11;

```
Qy         55 PIKTFRGAPPNSFEEFPPSALEGWT---------GATITVKIKCPEESASHLHVKNATMG 105
Db          4 PVRMFSLAP--------------WTPTPTPWLGGNTTAAAEAKC---------VFNEEFK 40

Qy        106 YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRS-ICTTVFYTNLAIADFLFCVT 164
Db         41 FI-------LLPISYGIVFVVGLPLNSWAMWIFVSRMRPWNATTTYMFNLAISDTLYVFS 93

Qy        165 LPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLP 224
Db         94 LPTLVYYYADRNNWPFGKVFCKIVRFLFYANLYSSILFLTCISVHRYMGICHPIRSLKWV 153

Qy        225 KHTYALVTCGLVWATVFLYMLP--FFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYF 282
Db        154 KTKHARLICVGVWLVVTICLIPNLIFVTTSS----KDNSTLCHD--TTKPEEFDHYVHYS 207

Qy        283 ISLAFFGFLIPFVLIIYCYAAIIRTL------------NAYDHRWLWYVKASLLILVIFT 330
Db        208 SSIMALLFGIPFLVIVVCYCLMAKRLCKRSFPSPSPRVPSYKKR---SIKMIIIVLTVFA 264

Qy        331 ICFAPSNIILIIHHANYYY----NNTDGLYFIYLIALCLGSLNSCLDPFLYFL 379
Db        265 ICFVPFHITRTLYYTSRYFQADCQTLNIIINFTYKITRPLASINSCLDPILYFM 317
```

RESULT 4
Q90X57
ID Q90X57 PRELIMINARY; PRT; 361 AA.
AC Q90X57;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE P2Y1 nucleotide receptor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC Xenopodinae; Xenopus.
```

```
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cheng A.W., Tsim K.W.;
RT   "Cloning of Xenopus P2Y1 Receptor.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF432354; AAL27614.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   361 AA;  41002 MW;  E5B2D605F5B57FED CRC64;

  Query Match             19.6%;  Score 422.5;  DB 13;  Length 361;
  Best Local Similarity   31.6%;  Pred. No. 2e-28;
  Matches  107;  Conservative   56;  Mismatches  137;  Indels   39;  Gaps    9;

Qy         95 SHLHVKNATMGYLTSSLSTK------LIPAIYLLVFVVGVPANAVTLWMLFPRTRSICT- 147
Db         15 STLLASGSSAGNVTKCSLTKTGFQFYYLPAVYIVVCITGFIGNSVAIWMFIFHMKPWSSI 74

Qy        148 TVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACIS 207
Db         75 SVYMFNLALADFLYVLSLPALIFYYFNKTDWIFGDALCKLQRFLFHVNLYGSILFLTCIS 134

Qy        208 INRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLP--FF----ILKQEYYLVQPDI 261
Db        135 VHRYTGVVHPLKSLGRLKKKNSIYISALVWFIVIAGISPILFFSGTGIRKNK-------T 187

Qy        262 TTCHDVHNTCESSSPFQLYYFI---SLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWY 318
Db        188 ITCFDT-----SSDEYLRSYFIYSMCTTVFGFCIPFILILGCYGLIVRALIYKDMNNAPL 242

Qy        319 VKAS----LLILVIFTICFAPSNIILIIH-------HANYYYNNTDGLYFIYLIALCLGS 367
Db        243 RKKSIYLVIIVLTVFAVSYLPFHVMKNLNLRARLDFQSPEMCNFNDRVYATYQVTRGLAS 302

Qy        368 LNSCLDPFLYFLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
Db        303 LNSCVDPILYFLAGDTFRRKLSRATRKASRRSEANVQSK 341


RESULT 5
Q99MT7
ID   Q99MT7      PRELIMINARY;      PRT;   359 AA.
AC   Q99MT7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G-protein coupled receptor GPR87.
GN   GPR87.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21172992; PubMed=11273702;
RA   Wittenberger T., Schaller H.C., Hellebrand S.;
RT   "An expressed sequence tag (est) data mining strategy succeeding in
RT   the discovery of new g-protein coupled receptors.";
RL   J. Mol. Biol. 307:799-813(2001).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF295366; AAK01866.1; -.
DR   MGD; MGI:1934133; Gpr87.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01655; UDPGLUCOSER.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   359 AA;  41462 MW;  2EB313C273EFCDEA CRC64;

  Query Match             17.9%;  Score 387;  DB 11;  Length 359;
  Best Local Similarity   28.2%;  Pred. No. 2.2e-25;
  Matches  100;  Conservative   70;  Mismatches  131;  Indels   54;  Gaps   11;

Qy         49 LAKPTLPIKTFRGAPPNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLT 108
Db          1 MAVPNVNVSTFAGN------------ELYSQASHTA------NSTSEGHGKNST---LH 38

Qy        109 SSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFY-TNLAIADFLFCVTLPF 167
Db         39 NKFDTIILPVLYLVIFVASILLNGLAVW-IFFHIRNKTSPIFYLKNIVVADLIMTLTFPF 97

Qy        168 KIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHT 227
Db         98 RIVRDAGFGPWYFEFILCRYTSVLFYANMYTSIVFLGLISVDRYLKVVKPFGDSRMYSIT 157

Qy        228 YALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAF 287
Db        158 FTKVLSVCVWVIMAILSLPNIILTNG----QP---TKENIHDCMKLKSPLGAKWHMAVTY 210

Qy        288 FG---FLIPFVLIIYCYAAIIRTLNAYDHRWL-------WYVKASLLILVIFTICFAPSN 337
Db        211 VDSCLFVAVLVILIGCYIAISRYIHKSSRQFISQSSRKRKHNQSIRVVVAVFFTCFLPYH 270

Qy        338 IILIIHHANYYYNNTDG---------LYFIYLIALCLGSLNSCLDPFLYFLMSKT 383
Db        271 LCRI----PFTFSNLDRLLDESAHKILYYCKEMTLFLSACNVCLDPIIYFFMCKS 321


RESULT 6
Q9BY21
ID   Q9BY21      PRELIMINARY;      PRT;   358 AA.
AC   Q9BY21;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Orphan G protein-coupled receptor 87.
GN   GPR87 OR GPR95.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21172992; PubMed=11273702;
RA   Wittenberger T., Schaller H.C., Hellebrand S.;
RT   "An expressed sequence tag (est) data mining strategy succeeding in
RT   the discovery of new g-protein coupled receptors.";
RL   J. Mol. Biol. 307:799-813(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21458557; PubMed=11574155;
RA   Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O.,
RA   Lewis T., Evans J.F., George S.R., O'Dowd B.F.;
RT   "Discovery and mapping of ten novel G protein-coupled receptor
RT   genes.";
RL   Gene 275:83-91(2001).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF237763; AAK01858.1; -.
DR   EMBL; AF411114; AAL26485.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01655; UDPGLUCOSER.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   358 AA;  41435 MW;  FDF157295BE5D10F CRC64;
```

```
  Query Match             17.1%;  Score 368.5;  DB 4;  Length 358;
  Best Local Similarity   29.3%;  Pred. No. 8.6e-24;
  Matches   89;  Conservative   64;  Mismatches  116;  Indels   35;  Gaps    9;

Qy        100 KNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFY-TNLAIAD 158
              || |   | :  | ::| :||::||  :  | : :| :||  |:  : :||  |: :||
Db         32 KNTT---LHNEFDTIVLPVLYLIIFVASILLNGLAVW-IFFHIRNKTSFIFYLKNIVVAD 87

Qy        159 FLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF 218
               :  :| ||:| :      | |  :||| |:|:|| ||| ||: |  |||:||| :| ||
Db         88 LIMTLTFPFRIVHDAGFGPWYFKFILCRYTSVLFYANMYTSIVFLGLISIDRYLKVVKPF 147

Qy        219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQ 278
               :  |:  |    ||  : :  ||  ||       ||   |  ::|:  :  ||
Db        148 GDSRMYSITFTKVLSVCVWVIMAVLSLPNIILTNG----QP---TEDNIHDCSKLKSPLG 200

Qy        279 LYYFISLAFFG---FLIPFVLIIYCYAAIIRTLNAYDHRWL-------WYVKASLLILVI 328
              : :  :: :     |:   |::| || || | ::    :::        : ::  ::: :
Db        201 VKWHTAVTYVNSCLFVAVLVILIGCYIAISRYIHKSSRQFISQSSRKRKHNQSIRVVVAV 260

Qy        329 FTICFAPSNIILIIHHANYYYNNTDG---------LYFIYLIALCLGSLNSCLDPFLYFL 379
              |  || | ::  |     : :::  |          ||:   | | | : | |||| :||
Db        261 FFTCFLPYHLCRI----PFTFSHLDRLLDESAQKILYYCKEITLFLSACNVCLDPIIYFF 316

Qy        380 MSKT 383
              | ::
Db        317 MCRS 320


RESULT 7
Q96JZ8
ID   Q96JZ8      PRELIMINARY;      PRT;   358 AA.
AC   Q96JZ8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CDNA FLJ14878 fis, clone PLACE1003238, weakly similar to probable G
DE   protein-coupled receptor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK027784; BAB55366.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01569; P2Y12PRNCPTR.
DR   PRINTS; PR01655; UDPGLUCOSER.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   358 AA;  41462 MW;  7822700C8E3CC7E9 CRC64;

  Query Match             17.1%;  Score 368.5;  DB 4;  Length 358;
  Best Local Similarity   29.3%;  Pred. No. 8.6e-24;
  Matches   89;  Conservative   64;  Mismatches  116;  Indels   35;  Gaps    9;

Qy        100 KNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFY-TNLAIAD 158
              || |   | :  | ::| :||::||  :  | : :| :||  |:  : :||  |: :||
Db         32 KNTT---LHNEFDTIVLPVLYLIIFVASILLNGLAVW-IFFHIRNKTSFIFYLKNIVVAD 87

Qy        159 FLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF 218
               :  :| ||:| :      | |  :||| |:|:|| ||| ||: |  |||:||| :| ||
Db         88 LIMTLTFPFRIVHDAGFGPWYFKFILCRYTSVLFYANMYTSIVFLGLISIDRYLKVVKPF 147

Qy        219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQ 278
               :  |:  |    ||  : :  ||  ||       ||   |  ::|:  :  ||
Db        148 GDSRMYSITFTKVLSVCVWVIMAVLSLPNIILTNG----QP---TEDNIHDCSKLKSPLG 200

Qy        279 LYYFISLAFFG---FLIPFVLIIYCYAAIIRTLNAYDHRWL-------WYVKASLLILVI 328
              : :  :: :     |:   |::| || || | ::    :::        : ::  ::: :
Db        201 VKWHTAVTYVNSCLFVAVLVILIGCYIAISRYIHKSSRQFISQSSRKRKHNQSIRVVVAV 260

Qy        329 FTICFAPSNIILIIHHANYYYNNTDG---------LYFIYLIALCLGSLNSCLDPFLYFL 379
              |  || | ::  |     : :::  |          ||:   | | | : | |||| :||
Db        261 FFTCFLPYHLCRI----PFTFSHLDRLLDESAQKILYYCKEITLFLSACNVCLDPIIYFF 316

Qy        380 MSKT 383
              | ::
Db        317 MCRS 320


RESULT 8
Q8R528
ID   Q8R528      PRELIMINARY;      PRT;   309 AA.
AC   Q8R528;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cysteinyl leukotriene 2 receptor.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ishii S.;
RT   "Mouse CysLT2 Gene.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB058930; BAB86881.1; -.
KW   Receptor.
SQ   SEQUENCE   309 AA;  35226 MW;  96FACC6B8AF96974 CRC64;

  Query Match             17.0%;  Score 367.5;  DB 11;  Length 309;
  Best Local Similarity   30.3%;  Pred. No. 9.2e-24;
  Matches   89;  Conservative   60;  Mismatches  114;  Indels   31;  Gaps   10;

Qy        100 KNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRT--RSICTTVFYTNLAIA 157
              :| |:     :   :  | |||::|  |    |  ::: :| :|  :|    ||  ||| :
Db         13 RNCTI----ENFKKEFYPIIYLIIFFWGALGNGFSIY-VFLQTCKKSTSVNVFMLNLATS 67

Qy        158 DFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHP 217
              ||||  ||||:  |:  |:||:||:: ||  :    | ||| ||  |  :|: |:|| |||
Db         68 DFLFISTLPFRADYYFRGSNWIFGDLACRVMSYSLYVNMYTSIYFLTVLSVVRFLATVHP 127

Qy        218 FTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPF 277
              |    :    | : ||::|  :        :  ||    : :| :| ::      |   |
Db        128 FRMFHVTSVRSAWILCGIIWVFIMASSALLLVNGQE---EKDNIISCLEL-----SPQKF 179

Qy        278 Q---LYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYD---------HRWLWYVKASLLI 325
              :   :   |::|  |||:||: :  ||  |||  |   :          ||    :   ::
Db        180 KSLLIMNHIAVA-VGFLLPFLTLTICYLLIIRILLKAEIPESGPRAAHRKA--LTTIVIA 236

Qy        326 LVIFTICFAPSNIILIIHHANYYYNNT-DGLYFIYLIALCLGSLNSCLDPFLYF 378
              :: | :|| | : :  :|   :  ::  | |:   :| | : : ||| :||||:
Db        237 MITFLLCFLPYHALRTLHLVTWDKDSCGDVLHKATVITLTMAAANSCFNPFLYY 290


RESULT 9
Q9PVY7
ID   Q9PVY7      PRELIMINARY;      PRT;   359 AA.
AC   Q9PVY7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Angiotensin receptor.
```

```
OS   Anguilla anguilla (European freshwater eel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC   Anguilla.
OX   NCBI_TaxID=7936;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tran Van Chuoi M., Dolphin C.T., Barker S., Clark A.J., Vinson G.P.;
RT   "Molecular cloning and characterization of the cDNA encoding the
RT   angiotensin II recpetor of european eel Anguilla anguilla.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AJ005132; CAB40835.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   359 AA;  39997 MW;  053726AEA827C271 CRC64;

  Query Match             17.0%;  Score 367.5;  DB 13;  Length 359;
  Best Local Similarity   27.0%;  Pred. No. 1.1e-23;
  Matches   99;  Conservative   61;  Mismatches  129;  Indels   77;  Gaps   10;

Qy         99 VKNATMG-----YLTSSLSTK------LIPAIYLLVFVVGVPANAVTLWMLF-FRTRSIC 146
              ::| |:|     ::| : | :      ||| :|   ||:|:  |:: : ::: :
Db          1 MENLTVGRTEGIHITCNTSGRHSYIYTLIPVVYGCNFVIGIVGNSMVVAVIYCYMKLKTV 60

Qy        147 TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACI 206
               :|  |||::|  | :|||    :   | || ||  ||:|:  :   |:| ||  | :
Db         61 ANIFVLNLAVSDLTFLITLPMWATFTAMGYNWPFGGFLCKASAGLTIFNLYTSIFFLTSL 120

Qy        207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHD 266
              ||:|||||||   |      || :|| |:||  ||  ||  : :  : :  |: | |
Db        121 SIDRYLAIVHPVRSRQRRTVVYARITCVLIWAFAFLLSLPTALSRDVFTINHPNTTVC-- 178

Qy        267 VHNTCESSSPFQLYYFISL--AFFGFLIPFVLIIYCYAAIIRTL------NAYDHRWLWY 318
               | :      :   | |  :  |||||||:|: ||  | | |       :   |
Db        179 --GTLDKHELSHVLVAIGLMKSVLGFLIPFVIIVTCYCLIGRALLEARRVQSSRSRGDEV 236

Qy        319 VKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIAL-------------- 363
              ::    ::: | :|: |  |                  :|::::||
Db        237 LQMLAAVVLAFFLCWVPHQI---------------FHFMHVLALLKVIENCPTLDIIDT 280

Qy        364 ------CLGSLNSCLDPFLYFLMSKT--RN-------------HSTAYLTKXXNDLREQ 401
                   |:   |||::| ||  : :   ||                ||   ||   : | :
Db        281 ALPFTICIAYFNSCMNPILYGFVGRNFRRNLLRLLRCGPGSAARHSHPSLTTKMSTLSYR 340

Qy        402 GQPSQR 407
                : |
Db        341 ASETLR 346


RESULT 10
Q9GK76
ID   Q9GK76      PRELIMINARY;      PRT;   342 AA.
AC   Q9GK76;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Platelet-activating factor receptor.
OS   Capra hircus (Goat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Capra.
OX   NCBI_TaxID=9925;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang W., Diehl J.R., Piumi F.;
RT   "Partial Sequence of Caprine Platelet-Activating Factor Receptor
RT   Gene.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF302764; AAG39982.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01559; DUFFYANTIGEN.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   342 AA;  39612 MW;  39D9379AED60FD4E CRC64;

  Query Match             16.6%;  Score 358.5;  DB 6;  Length 342;
  Best Local Similarity   30.7%;  Pred. No. 6e-23;
  Matches   99;  Conservative   48;  Mismatches  129;  Indels   47;  Gaps   12;

Qy        115 LIPAIYLLVFVVGVPANAVTLWM---LFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAY 171
              | |  | :|||:|| ||:  ||:   |:   :     :|  || :|| || ||||  | |
Db         17 LFPIFYSIVFVLGVIANSYVLWVFARLYPSKKFNEIKIFMVNLTMADLLFLVTLPLWIVY 76

Qy        172 HLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALV 231
              : |  :|:  : ||       |: | |||:  || |: ||: |:  |             :
Db         77 YYNQGDWILPKFLCNLAGCFFFINTYCSVAFLAVITYNRFQAVTRPIKTAQATTRKRGFL 136

Qy        232 TCGLVWATVFLYMLPFFIL----KQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAF 287
                ::| ::      ||:|     :       :|| | : :   : | |  : :   :
Db        137 LSLIIWVSIVGAASYFFVLDSTNSEPKKTGSGNITRCFEHYE--KGSIPVLIIHIFLV-- 192

Qy        288 FGFLIPFVLIIYCYAAIIRTL---------NA-YDHRWLWYVKASLLILVIFTICFAPSN 337
              | | : |::|::|   ||||          ||    | || |     :| :| ||| |
Db        193 FSFFLVFLIILFCNLVIIRTLLTQQVQMQRNAEVKRRALWMV---CTVLAVFVICFVP-- 247

Qy        338 IILIIHH--------ANYYYNNTD---GLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRN 385
                  ||        |   : :||   |:   : : ||| | |  ||| :| ||  | |
Db        248 -----HHLVQLPWTLAELGFQDTDFHQGINDAHQVTLCLLSTNCVLDPIIYCFLTKKFRK 302

Qy        386 HSTAYLTKXXNDLREQGQPSQRT 408
              |    ||:    :||  : |: |
Db        303 H----LTEKLYSMRESRKCSRAT 321


RESULT 11
Q9TTY5
ID   Q9TTY5      PRELIMINARY;      PRT;   342 AA.
AC   Q9TTY5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Platelet-activating factor receptor.
GN   PAFR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang W.S., Diehl J.R., Roudebush W.E.;
RT   "Partial Sequence of Bovine Platelet-Activating Factor Receptor
RT   Gene.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   He B., Tiemann U., Kanitz W., Weikard R., Laurent P., Schwerin M.,
RA   Schmidt P.;
RT   "Molecular characterization of bovine platelet-actvating factor
RT   receptor transcripts and their detection in different tissues of
RT   cattle.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AF187321; AAF01439.2; -.
DR   EMBL; AJ295321; CAC43290.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01559; DUFFYANTIGEN.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   342 AA;  39691 MW;  7C4236205AE937C9 CRC64;

  Query Match             16.5%;  Score 356.5;  DB 6;  Length 342;
  Best Local Similarity   30.7%;  Pred. No. 8.9e-23;
  Matches   99;  Conservative   48;  Mismatches  129;  Indels   47;  Gaps   12;

Qy        115 LIPAIYLLVFVVGVPANAVTLWM---LFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAY 171
              | |  | :|||:|| ||:  ||:   |:  :     :|  || :|| || |||| | |
Db         17 LFPIFYSIVFVLGVIANSYVLWVFARLYPSKKFNEIKIFMVNLTMADLLFLVTLPLWIVY  76

Qy        172 HLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALV 231
              : |  :|:  : ||      |: | |||:  || |: ||: |:  |          ::
Db         77 YYNQGDWILPKFLCNLAGCFFFINTYCSVAFLAVITYNRFQAVTRPIKTAQATTRKRGIL 136

Qy        232 TCGLVWATVFLYMLPFFIL----KQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAF 287
                ::| ::      ||:|    ::       :|| | : |     |      :: |:
Db        137 LSLIIWVSIVGAASYFFVLDSTNREPNKTGSANITRCFE-HYEKGSIPVLTIHIFL---V 192

Qy        288 FGFLIPFVLIIYCYAAIIRTL---------NA-YDHRWLWYVKASLLILVIFTICFAPSN 337
              | | : |::|::|   |||||          ||    | || |     :| :| ||| |
Db        193 FSFFLVFLIILFCNLVIIRTLLTQQVQIQRNAEVKRRALWMV---CTVLAVFIICFVP-- 247

Qy        338 IILIIHH--------ANYYYNNTD---GLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRN 385
                ||          |   : :||    :   : : ||| | |  ||| :| || | |
Db        248 -----HHLVQLPWTLAELGFQDTDFHQAINDAHQVTLCLLSTNCVLDPIIYCFLTKKFRK 302

Qy        386 HSTAYLTKXXNDLREQGQPSQRT 408
              |    ||:    :||  : |: |
Db        303 H----LTEKLYSMRESRKCSRAT 321


RESULT 12
Q9R1V0
ID   Q9R1V0      PRELIMINARY;      PRT;   367 AA.
AC   Q9R1V0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CC chemokine LARC specific receptor.
GN   CMKBR6 OR MCCR6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanaka Y.;
RT   "Molecular Cloning of Murine Homologue of CCR6, the Specific Receptor
RT   for CC Chemokine LARC.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB016031; BAA82443.1; -.
DR   HSSP; P34996; 1DDD.
DR   MGD; MGI:1333797; Cmkbr6.
DR   InterPro; IPR004067; CC_chemkine6.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01529; CHEMOKINER6.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   367 AA;  42082 MW;  207FD98B2F7A6DD3 CRC64;

  Query Match             16.4%;  Score 353;  DB 11;  Length 367;
  Best Local Similarity   31.0%;  Pred. No. 1.9e-22;
  Matches   96;  Conservative   49;  Mismatches  121;  Indels   44;  Gaps   12;

Qy        113 TKL-IPAIYLLVFVVGVPAN-AVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIA 170
              ||: :|  | |: | |:  | ' :   |::     | |:  |:|| | || :||||
Db         37 TKVFVPIAYSLICVFGLLGNIMVVMTFAFYKKARSMTDVYLLNMAITDILFVLTLPFWAV  96

Qy        171 YHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVH---PFTYRGLPKHT 227
              |   | |||  : ||:      :  |  | :|||||||::||:|||     |  |     |
Db         97 THAT-NTWVFSDALCKLMKGTYAVNFNCGMLLLACISMDRYIAIVQATKSFRVRSRTL-T 154

Qy        228 YALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAF 287
              :: | |  ||    :   | ||  ::| |   |:  |   : :       ::|      |
Db        155 HSKVICVAVWFISIIISSPTFIFNKKYELQDRDV--CEPRYRSVSEPITWKLLGLGLELF 212

Qy        288 FGFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNIILIIH 343
              |||  | : :::||  ||:||    |:  ||    ::  : ::::|  |  | |::|::
Db        213 FGFFTPLLFMVFCYLFIIKTLVQAQNSKRHR---AIRVVIAVVLVFLACQIPHNMVLLVT 269

Qy        344 HAN-------------YYYNNTDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNHST 388
                |               |  |       :|  |  |:|||:| || |:  | ||
Db        270 AVNTGKVARSCSTEKVLAYTRN---------VAEVLAFLHSCLNPVLYAFIGQKFRN--- 317

Qy        389 AYLTKXXNDL 398
              |  |   |:
Db        318 -YFMKIMKDV 326


RESULT 13
Q96P68
ID   Q96P68      PRELIMINARY;      PRT;   337 AA.
AC   Q96P68;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor (Putative G-protein coupled receptor).
GN   GPR80 OR GPCR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21458557; PubMed=11574155;
RA   Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O.,
RA   Lewis T., Evans J.F., George S.R., O'Dowd B.F.;
RT   "Discovery and mapping of ten novel G protein-coupled receptor
RT   genes.";
RL   Gene 275:83-91(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT   "Identification of G protein-coupled receptor genes from the human
RT   genome sequence.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF411109; AAL26480.1; -.
DR   EMBL; AB083598; BAB89311.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   337 AA;  38251 MW;  6814EA0044756CE6 CRC64;

  Query Match             16.3%;  Score 351;  DB 4;  Length 337;
  Best Local Similarity   29.1%;  Pred. No. 2.6e-22;
  Matches   87;  Conservative   62;  Mismatches  120;  Indels   30;  Gaps    8;

Qy        102 ATMGYLTSS---LSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRS-ICTTVFYTNLAIA 157
              |  |   |     |    :| || ::|:|| | ||| :    |: |    :|:   |||
Db         19 AAFGNCTDENIPLKMHYLPVIYGIIFLVGFPGNAVVISTYIFKMRPWKSSTIIMLNLACT  78

Qy        158 DFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHP 217
              | |:  :||| | |: :| ||:||: :|:     |: |:| ||| | | || ||  |:||
```

```
Db         79 DLLYLTSLPFLIHYYASGENWIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRYCVIIHP 138

Qy        218 FTYRGLPKHTYALVTCGLVWATVFLYMLP--FFILKQEYYLVQPDITTCHDVHNTCESSS 275

Db        139 MSCFSIHKTRCAVVACAVVWIISLVAVIPMTFLITSTN----RTNRSACLDLTSSDELNT 194

Qy        276 PFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLN---------AYDHRWLWYVKASLLIL 326

Db        195 --IKWYNLILTATTFCLPLVIVTLCYTTIIHTLTHGLQTDSCLKQKAR-----RLTILLL 247

Qy        327 VIFTICFAPSNIILIIHHANYYYNNT----DGLYFIYLIALCLGSLNSCLDPFLYFLMS 381

Db        248 LAFYVCFLPFHILRVIRIESRLLSISCSIENQIHEAYIVSRPLAALNTFGNLLLYVVVS 306


RESULT 14
O46685
ID   O46685      PRELIMINARY;      PRT;   361 AA.
AC   O46685;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Orphan G protein-coupled receptor BRGR1B.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Ferrer-Martinez A., Felipe A., Casado J., Pastor-Anglada M.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; U88367; AAC05612.1; -.
DR   EMBL; U88366; AAC05611.1; -.
DR   HSSP; P02699; 1F88.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01564; OGR1RECEPTOR.
DR   PRINTS; PR01649; PSYCHOSINER.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ   SEQUENCE   361 AA;  40905 MW;  4802E77C9A45DEE1 CRC64;

  Query Match             16.1%;  Score 347;  DB 6;  Length 361;
  Best Local Similarity   27.4%;  Pred. No. 6.1e-22;
  Matches   86;  Conservative   61;  Mismatches  127;  Indels   40;  Gaps    8;

Qy        104 MGYLTS-----------SLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTT-VFY 151

Db          1 MGNITADNTSMNCDIDHTIHQTLAPVVYVMVLVVGFPANCLSLYYGYLQIKARNELGVYL 60

Qy        152 TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRY 211

Db         61 CNLTVADLFYICSLPFWLQYVLQHDHWSHDDLSCQVCGILLYENIYISVGFLCCISIDRY 120

Qy        212 LAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTC 271

Db        121 LAVAHPFRFHQFRTLKAAMGVSALIWVKELLTSI--------YFLMHEEVVEDADRHRVC 172

Qy        272 ESSSPFQ-------LYYFISLAFFGFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVK 320

Db        173 FEHYPLEPRQRGINYYRFL----VGFLFPICLLLASYRGILRAVRRSHGTQKSRKDQIQR 228

Qy        321 ASLLILVIFTICFAPSNIILIIHHANYYYNNTD---GLYFIYLIALCLGSLNSCLDPFLY 377

Db        229 LVLSTVVIFLACFLPYHVLLLVR--SLWESSCDFAKGIFNAYHFSLLLTSFNCVADPVLY 286

Qy        378 FLMSKTRNHSTAYL 391

                :|:| :   | |
Db        287 CFVSETTHRDLARL 300


RESULT 15
Q9EPP3
ID   Q9EPP3      PRELIMINARY;      PRT;   359 AA.
AC   Q9EPP3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Angiotensin II type 1 receptor.
GN   AT1.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=GALL BLADDER;
RA   Busmann A., Noack T., Zucht HD., Forssmann W.G., Meyer M.;
RT   "Isolation of Angiotensin I from human hemofiltrate as a potent
RT   modulator of KATP channels.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AJ301623; CAC21550.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   359 AA;  41045 MW;  D00F4C2C472DE5CC CRC64;

  Query Match             16.0%;  Score 345;  DB 11;  Length 359;
  Best Local Similarity   27.6%;  Pred. No. 9e-22;
  Matches   92;  Conservative   58;  Mismatches  129;  Indels   54;  Gaps    9;

Qy         89 CPEESA-SHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTR-SIC 146

Db         18 CPKDGRHSYIFV----------------MIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTV 62

Qy        147 TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACI 206

Db         63 ASVFLLNLALADICFLLTLPLWAVYTAMEYRWPFGNYMCKIASASVSFNLYASVFLLTCL 122

Qy        207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHD 266

Db        123 SIDRYLAIVHPMKSRLRRTMLVAKVTCVIIWLMAGLASLPAVIHRNVFFIENTNITVCAF 182

Qy        267 VHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTL-NAYD-----HRWLWYVK 320

Db        183 HYESQNSTLPIGL--GLTKNILGFMFPFLIILTSYTLIWKALKKAYEIQKNKPRNDDIFK 240

Qy        321 ASLLILVIFTICFAPSNIIL---------IIHHANYYYNNTDGLYFIYLIALCLGSLNSC 371

Db        241 IIMAIVLFFFFSWVPHQIFTFLDVLIQLGIIHDCKI----SDIVDTAMPITICIAYFNNC 296

Qy        372 LDPFLY----------------FLMSKTRNHST 388

Db        297 LNPLFYGFLGKKFKKYFLQLLKYIPPKAKSHST 329


RESULT 16
Q9TTY6
ID   Q9TTY6      PRELIMINARY;      PRT;   296 AA.
AC   Q9TTY6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Platelet-activating factor receptor (Fragment).
```

```
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang W.S., Diehl J.R., Murphy K.E.;
RT   "Partial Sequence of Canine Platelet-Activating Factor Receptor
RT   Gene.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF186831; AAF01435.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     296    296
SQ   SEQUENCE   296 AA;  33902 MW;  5494C0771C45E101 CRC64;

  Query Match             15.8%;  Score 341.5;  DB 6;  Length 296;
  Best Local Similarity   30.5%;  Pred. No. 1.5e-21;
  Matches   91;  Conservative   43;  Mismatches  127;  Indels   37;  Gaps   10;

Qy        115 LIPAIYLLVFVVGVPANAVTLWM---LFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAY 171
              | | :| ::||:|| ||:  ||:   |:   :     :|  || :|| || |||| | |
Db          9 LFPIVYSIIFVLGVVANSYVLWVFACLYPSKKLNEIKIFMVNLTMADLLFLVTLPLWIIY 68

Qy        172 HLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALV 231
              : |  ||:    ||      |: | |||:  || |: ||: |:  |          :
Db         69 YHNQGNWILPSFLCNLAGCFFFINTYCSVAFLAVITYNRFQAVTRPIKTAQATTRKRGFL 128

Qy        232 TCGLVWATVFLYMLPFFILKQEYYLVQP----DITTCHDVHNTCESSSPFQLYYFISLAF 287
                ::|  :      ||||     :       :|| | : |     |      :: ||
Db        129 VSLVIWVAIVAAASYFFILDSTNVVPSKSGSGNITRCFE-HYEKGSVPVLIVHVFI---V 184

Qy        288 FGFLIPFVLIIYCYAAIIRTL---------NA-YDHRWLWYVKASLLILVIFTICFAPSN 337
              ||  | |::| :|   ||| |          ||    | || |     :| :| :|| | :
Db        185 FGLFIVFLIIFFCNVVIIRRLLMQPVQLQHNAEVKRRALWMV---CTVLAVFIVCFVPHH 241

Qy        338 II--------LIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNH 386
              ::        |    ::::     |:   : : ||| | |  ||| :| ||  | | |
Db        242 LVQLPWTLAELGFQSSSFH----QGINDAHQVTLCLLSTNCVLDPIIYCFLTKKFRKH 295


RESULT 17
Q9ERK9
ID   Q9ERK9      PRELIMINARY;      PRT;   328 AA.
AC   Q9ERK9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   P2Y6 receptor (Hypothetical 36.7 kDa protein).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVEV;
RX   MEDLINE=21160052; PubMed=11259526;
RA   Lazarowski E.R., Rochelle L.G., O'Neal W.K., Ribeiro C.M.P.,
RA   Grubb B.R., Zhang V., Harden T.K., Boucher R.C.;
RT   "Cloning and functional characterization of two murine uridine
RT   nucleotide receptors reveal a potential target for correcting ion
RT   transport deficiency in cystic fibrosis gallbladder.";
RL   J. Pharmacol. Exp. Ther. 297:43-49(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF298899; AAG24619.1; -.
DR   EMBL; BC027331; AAH27331.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Hypothetical protein; Receptor.
SQ   SEQUENCE   328 AA;  36721 MW;  00F9DF5ADADF903E CRC64;

  Query Match             15.8%;  Score 340;  DB 11;  Length 328;
  Best Local Similarity   30.7%;  Pred. No. 2.2e-21;
  Matches   90;  Conservative   47;  Mismatches  122;  Indels   34;  Gaps    9;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFYT-NLAIADFLFCVTLPFKIAYHL 173
              |:  :| :| |||:| |    :  :     |::  :  || |||:|| ::  :||  | :
Db         28 LLTPVYSVVLVVGLPLNICVIAQICASRRTLTRSAVYTLNLALADLMYACSLPLLIYNYA 87

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF-TYRGLPKHTYALVT 232
              |::| ||:: ||    :|| |:: ||| | |||  ||| | ||  ::        | |
Db         88 RGDHWPFGDLACRFVRFLFYANLHGSILFLTCISFQRYLGICHPLASWHKRGGRRAAWVV 147

Qy        233 CGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPF----QLYYFISLAFF 288
              ||:||  |    ||  :       :| : |:|:|:      | |      | | ::|
Db        148 CGVVWLAVTAQCLPTAVFAATG--IQRNRTVCYDL------SPPILSTRYLPYGMALTVI 199

Qy        289 GFLIPFVLIIYCYAAIIRTLNAYD--------HRWLWYVKASLLILVIFTICFAPSNIIL 340
              |||:||: :: ||  : | |   |          |     : ::::  :| | | | :|
Db        200 GFLLPFIALLACYCRMARRLCRQDGPAGPVAQERRSKAARMAVVVAAVFAISFLPFHITK 259

Qy        341 IIHHANYYYNNTDGLYFIYLIALC--------LGSLNSCLDPFL-YFLMSKTR 384
                : |     :| |:    |              |:|| ||| | ||   | |
Db        260 TAYLA---VRSTPGVSCPVLETFAAAYKGTRPFASVNSVLDPILFYFTQQKFR 309


RESULT 18
Q9N0U1
ID   Q9N0U1      PRELIMINARY;      PRT;   359 AA.
AC   Q9N0U1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Angiotensin II type 1 receptor.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ADRENAL CORTEX;
RA   Bird I.M., Millican D.S.;
RT   "Isolation of a Full Length Ovine Angiotensin II Type-1 Receptor (AT1-
RT   R) cDNA.";
RL   Endocr. Res. 0:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF254119; AAF66063.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   359 AA;  41062 MW;  C07010EDB81110EB CRC64;

  Query Match             15.8%;  Score 340;  DB 6;  Length 359;
  Best Local Similarity   30.7%;  Pred. No. 2.4e-21;
  Matches   89;  Conservative   47;  Mismatches  120;  Indels   34;  Gaps    6;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTR-SICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              :|| :| ::|||:  |:: : :::|  :     :||  |||:||  | :|||     |
Db         30 MIPTLYSIIFVVGLFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLTLPLWAVYTA 89
```

```
Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
              |  ||  ||:  :     |:| |: || |:||:|||||||   |       | |||
Db         90 MEYRWPFGNYLCKIASGSVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTC 149

Qy        234 GLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIP 293
              ::|    |  ||  | :  :::   :|| |   : :  |: |  |   ::    ||| |
Db        150 IIIWLLAGLASLPTIIHRNVFFIENTNITVCAFHYESQNSTLPVGL--GLTKNILGFLFP 207

Qy        294 FVLIIYCYAAIIRTL-NAYD-----HRWLWYVKASLLILVIFTICFAPSNIILIIHHANY 347
              |::|:  |  | :||  ||:      |     |  | |:: |   : |  |   :
Db        208 FLIILTSYTLIWKTLKKAYEIQKNKPRKDDIFKIILAIVLFFFFSWVPHQIFTFM----- 262

Qy        348 YYNNTDGLYFIYL--------------IALCLGSLNSCLDPFLYFLMSK 382
                  | |  : |               | :||   |:||:|  |  : |
Db        263 -----DVLIQLGLIRDCKIEDIVDTAMPITICLAYFNNCLNPLFYGFLGK 307


RESULT 19
O97724
ID   O97724         PRELIMINARY;   PRT;   334 AA.
AC   O97724;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CCR5 chemokine receptor (Chemokine receptor 5).
GN   CCR5.
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kovacs E.M., Baxter G.D., Robinson W.F.;
RT   "Feline peripheral blood mononuclear cells express message for both
RT   CXC and CC type chemokine receptors.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-73 FROM N.A.
RA   Leutenegger C.M., Mislin C., Lutz H.;
RT   "Molecular cloning and expression of the 5' region of the feline
RT   chemokine receptor (CCR) 5.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ009815; CAA08838.1; -.
DR   EMBL; AF059714; AAC69545.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01559; DUFFYANTIGEN.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   334 AA;  37594 MW;  695B0FE18DB5532E CRC64;

  Query Match             15.7%;  Score 338;  DB 6;  Length 334;
  Best Local Similarity   28.2%;  Pred. No. 3.4e-21;
  Matches   84;  Conservative   66;  Mismatches  118;  Indels   30;  Gaps    9;

Qy        111 LSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNLAIADFLFCVTLPFK 168
              :: :|:| :| |||: |   | : : :|    : | : | |:  ||||:| ||  ||||
Db         28 IAARLLPPLYSLVFLSGFVGNLLVVLILINCKKLRGM-TDVYLLNLAISDLLFLFTLPFW 86

Qy        169 IAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTY 228
              .  |. ||  ||||:|:|:  | :::   :    :  ::::|||||||         |:
Db         87 AHYAANG--WVFGDVMCKTVTGLYHVGYFGGNFFIILLTVDRYLAIVHAVFALKARTVTF 144

Qy        229 ALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSP--FQLYYFISLA 286
              ||  : ||   :  ||  :  :  :          |   :|  | |  :  :: :: :
Db        145 GAVTSAVTWAAAVVTSLPGCVFNKLQW---------EDSFYSCRPSFPPGWNTFHAVTRS 195

Qy        287 FFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKASLLILVIFTICFAPSNIILI 341
              | ::| :::| ||:||:|||         ||    ||   :|:::: : :||:||:|:
Db        196 VLGLVLPLLVMIVCYSAILRTLFRCRNEKKKHR---AVKLIFVIMIVYFLFWAPNNIVLL 252

Qy        342 IHHANYYYN-----NTDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNHSTAYLTK 393
              :      :|     :|  |      :   ||  : |::| :| |:  | | | : :  |
Db        253 LSTFPESFNVSDCQSTSQLDQAMQVTETLGMTHCCINPIIYAFVGEKFRRHLSVFFRK 310


RESULT 20
Q8VE54
ID   Q8VE54      PRELIMINARY;      PRT;   298 AA.
AC   Q8VE54;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to G protein-coupled receptor 23 (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC019743; AAH19743.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   298 AA;  33708 MW;  A09209FBABCE65F2 CRC64;

  Query Match             15.6%;  Score 337.5;  DB 11;  Length 298;
  Best Local Similarity   31.3%;  Pred. No. 3.3e-21;
  Matches   78;  Conservative   45;  Mismatches   99;  Indels   27;  Gaps    6;

Qy        147 TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACI 206
              | :| ||||::| ||  ||||||  |: |  :| ||: ||: :   |  |:| |:| | ||
Db          2 TAIFITNLALSDLLFVCTLPFKIFYNFN-RHWPFGDTLCKISGTAFLTNIYGSMLFLTCI 60

Qy        207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHD 266
              |::|:||||:||  | :     : : |  ||  |    :   :             |  ::
Db         61 SVDRFLAIVYPFRSRTIRTRRNSAIVCAGVWILVLSGGISASLF---------STTNVNN 111

Qy        267 VHNTCESSSPFQLY--YFISLAFF----GFLIPFVLIIYCYAAIIR------TLNAYDHR 314
                ||      :::  |  :  |    ||:|| :| : | : ::|      ||:
Db        112 ATTTCFEGFSKRVWKTYLSKITIFIEVVGFIIPLILNVSCSSVVLRTLRKPATLSQIGTN 171

Qy        315 WLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLY-----FIYLIALCLGSLN 369
                :|   : : :| :|| | | :| ::        |: |       :| | ||| :||
Db        172 KKKVLKMITVHMAVFVVCFVPYNSVLFLYALVRSQAITNCLLERFAKIMYPITLCLATLN 231

Qy        370 SCLDPFLYF 378
                | |||:|:
Db        232 CCFDPFIYY 240


RESULT 21
Q8QGQ4
ID   Q8QGQ4      PRELIMINARY;      PRT;   390 AA.
AC   Q8QGQ4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Type five-like somatostatin receptor.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lin X., Peter R.E.;
```

```
RT   "Cloning and expression of a type five-like somatostatin receptor in
RT   goldfish.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF252879; AAM18805.1; -.
KW   Receptor.
SQ   SEQUENCE   390 AA;  44136 MW;  EC12119A4B6CF9A8 CRC64;

  Query Match             15.6%;  Score 337;  DB 13;  Length 390;
  Best Local Similarity   31.9%;  Pred. No. 4.7e-21;
  Matches   94;  Conservative   51;  Mismatches  102;  Indels   48;  Gaps   10;

Qy        112 STKLIPAIYLLVFVVGVPANAVTLWMLFFRTR-SICTTVFYTNLAIADFLFCVTLPFKIA 170
              |||::  |||:|||||:  |:: ::::   |:    | ::  |||:|| |: : |||  |
Db         39 STKILAVIYLVVFVVGLTGNSLAIFVVLRYTKMKTATNMYILNLAVADELYILGLPFLTA 98

Qy        171 YHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYAL 230
              ::: |  | ||  |||        : : |   |  :||:||:|:|||       :   |
Db         99 HNMLG-YWPFGNFLCRILMWTDSISQFTSTFCLTVMSIDRYMAVVHPIRSARWRRPRVAK 157

Qy        231 VTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCH-------DVHNTCESSSPFQLYYFI 283
              |  :|||   |  ||  |    |  |||:: ||:        || :|      | ||
Db        158 VINSMVWALSCLLTLPVII----YCDVQPELNTCNLSWPEPRDVWSTA-----FILY--- 205

Qy        284 SLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWYVKASL---------------LILVI 328
                |  ||  | ::|  ||  |:  : :   |      | |                :|:|:
Db        206 -TAMLGFFFPLMVICLCYLLIVIKVKSASAR------AGLSKRGRSEKKVTRMVVIIVVV 258

Qy        329 FTICFAPSNIILIIHHANYYYNNT--DGLYFIYLIALCLGSLNSCLDPFLYFLMS 381
              | :|: |  || |::  :    |:  |:||: :|   |  :||| :| ||  :|
Db        259 FVLCWLPFFIINILNLISTLPENSLMTGIYFLTVI---LTYVNSCANPLLYSFLS 310


RESULT 22
Q9XSD4
ID   Q9XSD4      PRELIMINARY;      PRT;   342 AA.
AC   Q9XSD4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Platelet-activating factor receptor.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Diehl J.R., Murphy K.E., Roudebush W.E.;
RT   "Porcine (Sus scrofa) platelet-activating factor receptor DNA.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF124054; AAD28739.2; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01559; DUFFYANTIGEN.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   342 AA;  39435 MW;  830EA7542B528DAC CRC64;

  Query Match             15.6%;  Score 336.5;  DB 6;  Length 342;
  Best Local Similarity   29.9%;  Pred. No. 4.6e-21;
  Matches   89;  Conservative   47;  Mismatches  125;  Indels   37;  Gaps   11;

Qy        115 LIPAIYLLVFVVGVPANAVTLWML--FFRTRSIC-TTVFYTNLAIADFLFCVTLPFKIAY 171
              | |  | ::||:|| ||:  ||:    : :: :    :|  || :|| || |||| | |
Db         17 LFPIFYSIIFVLGVIANSYVLWVFARVYPSKKLNEIKIFMLNLTMADLLFLVTLPLWIIY 76

Qy        172 HLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALV 231
              : :  ||:  : ||      |: | |||:  || |: ||: |:  |          :
Db         77 YYHEGNWILPKFLCNLAGCFFFINTYCSVSFLAVITYNRFQAVTRPIKTAQATTRKRGIS 136

Qy        232 TCGLVWATVFLYMLPFFILKQEYYLVQP----DITTCHDVHNTCESSSPFQLYYFISLAF 287
                ::|  :      ||:|      :      ::| | : :   : | |  : :   :

Db        137 LSLIIWVAMVAAASYFFVLDSTNIELSKTGAGNLTRCFEHYE--KGSMPVLIIHIFLV-- 192

Qy        288 FGFLIPFVLIIYCYAAIIRTL---------NA-YDHRWLWYVKASLLILVIFTICFAPSN 337
              | | : |::|::|   |||||         ||    | || |     :| :| ||| | :
Db        193 FSFFLVFLVILFCNLVIIRTLLTQSVQMQRNAEVKRRALWMV---CTVLAVFIICFVPHH 249

Qy        338 II--------LIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNH 386
              |:        |     |::    |     : | ||| | |  ||| :| ||  | | |
Db        250 IVQLPWTLAELGFQSGNFHQAINDA----HQITLCLLSTNCVLDPIIYCFLTKKFRKH 303


RESULT 23
Q9GLN9
ID   Q9GLN9      PRELIMINARY;      PRT;   359 AA.
AC   Q9GLN9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Angiotensin II type-1 receptor.
GN   AGTR1.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20469400; PubMed=11013071;
RA   Dufour C., Casane D., Denton D., Wickings J., Corvol P.,
RA   Jeunemaitre X.;
RT   "Human-Chimpanzee DNA sequence variation in the four major genes of
RT   the renin angiotensin system.";
RL   Genomics 69:14-26(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF193445; AAG28410.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   359 AA;  41033 MW;  266133A7A3E911A6 CRC64;

  Query Match             15.5%;  Score 334;  DB 6;  Length 359;
  Best Local Similarity   29.3%;  Pred. No. 7.9e-21;
  Matches   85;  Conservative   50;  Mismatches  121;  Indels   34;  Gaps    6;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTR-SICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              :|| :| ::|||:  |:: : :::|  :     :||  |||:||  | :|||    |
Db         30 MIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLTLPLWAVYTA 89

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
                | ||  ||:  :      |:| |: || |:||:|||||||   |       | |||
Db         90 MEYRWPFGNYLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTC 149

Qy        234 GLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIP 293
              ::|    |  ||  | :  :::   :|| |   : :  |: | |  ::    ||| |
Db        150 IIIWLLAGLASLPAIIHRNVFFIENTNITVCAFHYESQNSTLPIGL--GLTKNILGFLFP 207

Qy        294 FVLIIYCYAAIIRTL-NAYD-----HRWLWYVKASLLILVIFTICFAPSNIILIIHHANY 347
              |::|:  |  | : |  ||:       |     |  : |::  |   : | |   :
Db        208 FLIILTSYTLIWKALKKAYEIQKNKPRNDDIFKIIMAIVLFFFFSWIPHQIFTFL----- 262

Qy        348 YYNNTDGLYFIYL--------------IALCLGSLNSCLDPFLYFLMSK 382
                | |  : :                | :|:   |:||:|  |  : |
Db        263 -----DVLIQLGIIRDCRIADIVDTAMPITICIAYFNNCLNPLFYGFLGK 307


RESULT 24
Q9TUS6
ID   Q9TUS6      PRELIMINARY;      PRT;   339 AA.
AC   Q9TUS6;
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   C-C chemokine receptor 5 (Fragment).
GN   CCR5.
OS   Papio cynocephalus papio.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=100937;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kunstman K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA   Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinksy S.;
RT   "Sequences of the CCR5 genes from diverse simian and prosimian
RT   species.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF161996; AAD47751.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  39032 MW;  327E49A1B14D38B9 CRC64;

  Query Match             15.5%;  Score 333.5;  DB 6;  Length 339;
  Best Local Similarity   25.9%;  Pred. No. 8.3e-21;
  Matches   82;  Conservative   69;  Mismatches  122;  Indels   43;  Gaps   10;

Qy        106 YLTS---------SLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNL 154
                  | ||          :: :|:| :| |||: |  | : : :|    | :|: | ::  ||
Db          7 YYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLLNL 65

Qy        155 AIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAI 214
                  ||:| || :|:||    |      | ||  :|:  | :::   :  |  :  ::|:|||||
Db         66 AISDLLFLLTVPFWA--HYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAI 123

Qy        215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESS 274
                  ||          |: :||  : |       ||  |  :         :    :|||| |
Db        124 VHAVFALKTKTVTFGVVTSVITWVVAVFASLPGIIFTR---------SQREGLHNTCSSH 174

Qy        275 SPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKASL 323
                |:  | |      : :   | ::| :::: ||: |::||        ||    |:
Db        175 FPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRLIF 231

Qy        324 LILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFLY- 377
                |:::: : :|| ||:|:::    ::     :::: |     :   ||  : |::| :|
Db        232 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYA 291

Qy        378 FLMSKTRNHSTAYLTK 393
                |:  | ||:   :  |
Db        292 FVGEKFRNYLLVFFQK 307


RESULT 25
Q8QFR5
ID   Q8QFR5      PRELIMINARY;      PRT;   353 AA.
AC   Q8QFR5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CXCR2 like protein.
GN   CXCR2.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Moepps B.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Moepps B.;
RL   Thesis (2000), Department of Pharmacology and Toxicology,
RL   University of Ulm, Ulm, Germany.
DR   EMBL; AJ312936; CAC85089.1; -.
SQ   SEQUENCE   353 AA;  40156 MW;  143E7E8C5B8B2F2B CRC64;

  Query Match             15.4%;  Score 333;  DB 13;  Length 353;
  Best Local Similarity   27.2%;  Pred. No. 9.5e-21;
  Matches   93;  Conservative   61;  Mismatches  128;  Indels   60;  Gaps   13;

Qy         67 FEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVV 126
                | :  || :   ||    :   |                     |  ::   :  :| ||| :
Db         10 FNDIDFSDIP--TGFPTVISAPCKS---------------TWVINKYFVVVVYALVFFL 51

Qy        127 GVPANAVTLWMLF-FRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLC 185
                 |  |:: : :::  : :   | |:  :||||| ||  ||||  ||    : ||||  :|
Db         52 NVVGNSLVVLVIYNNKLKRSSTDVYLLHLAIADLLFATTLPFWAAY--KASQWVFGIFMC 109

Qy        186 RATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYML 245
                :| :|:   | |  |||||||::|||||||         :|    :  | :|    :  |
Db        110 KAVSVLQEVNFYSGILLLACISVDRYLAIVHATEAVTQKRHWVKFICLG-IWIFSLVVSL 168

Qy        246 PFFILKQEYYLVQPDITTCHD-VHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAI 304
                |  : :  :   : |  ||| : |  |::  : :   |     || || :::::||
Db        169 PTLLFRTVFKSPR-DAYVCHDSIGN--ENTEDWMIILRIGRHLVGFFIPLLIMLFCYGFT 225

Qy        305 IRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNIILII----------------HH 344
                |:||    ::  ||    :|    ::: | ||: | |: :|:                |
Db        226 IKTLYQTKSSQKHR---AMKVIFAVVLAFLICWLPYNLTVIVDSLMRTRFINETCEKREH 282

Qy        345 ANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRN 385
                :   : |:             |  :||::| || |:  |  |
Db        283 LDAALSTTE----------IFGYTHSCINPILYAFIGQKFWN 314


RESULT 26
Q9UNW8
ID   Q9UNW8      PRELIMINARY;      PRT;   380 AA.
AC   Q9UNW8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor.
GN   G2A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Weng Z., Witte O.N.;
RT   "A DNA damage and stress inducible G protein-coupled receptor blocks
RT   cells in G2/M.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF083955; AAD47380.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01563; G2ARECEPTOR.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   380 AA;  42499 MW;  6DE63D17275ECD74 CRC64;

  Query Match             15.4%;  Score 332;  DB 4;  Length 380;
  Best Local Similarity   26.3%;  Pred. No. 1.2e-20;
  Matches   91;  Conservative   66;  Mismatches  153;  Indels   36;  Gaps    9;

Qy         71 PFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPA 130
```

```
Db          3 PMLLKNGYNGNATPVTTTAPWAS---LGLSAKTCNNVSFBESRIVLVVVYSAVCTLGVPA 59

Qy        131 NAVTLWMLFFRT-RSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATT 189

Db         60 NCLTAWLALLQVLQGNVLAVYLLCLALCELLYTGTLPLWVIYIRNQHRWTLGLLACKVTA 119

Qy        190 VIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFI 249

Db        120 YIFFCNIYVSILFLCCISCDRFVAVVYALESRGRRRRRTAILISACIFILVGIVHYPVFQ 179

Qy        250 LKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRT-- 307

Db        180 TE--------DKETCFDM---LQMDSRIAGYYYARFT-VGFAIPLSIIAFTNHRIFRSIK 227

Qy        308 ----LNAYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTD---------G 354

Db        228 QSMGLSAAQKAKVKH--SAIAVVVIFLVCFAPYHLVLLVKAAAFSYYRGDRNAMCGLEER 285

Qy        355 LYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXNDLRE 400

Db        286 LYTASVVFLCLSTVNGVADPIIYVLAT---DHSRQEVSRIHKGWKE 328


RESULT 27
Q9TV48
ID   Q9TV48      PRELIMINARY;      PRT;   352 AA.
AC   Q9TV48;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CC chemokine receptor type 5.
GN   CCR5.
OS   Cercopithecus ascanius (black-cheeked white-nosed monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=36223;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=52;
RX   MEDLINE=99335215; PubMed=10408730;
RA   Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
RA   Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT   "Mutations in CCR5-coding sequences are not associated with SIV
RT   carrier status in African nonhuman primates.";
RL   AIDS Res. Hum. Retroviruses 15:931-939(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=52;
RA   Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
RA   Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF035216; AAD44009.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   352 AA;  40650 MW;  7906256AA3945266 CRC64;

  Query Match             15.4%;  Score 331.5;  DB 6;  Length 352;
  Best Local Similarity   26.4%;  Pred. No. 1.3e-20;
  Matches   84;  Conservative   67;  Mismatches  120;  Indels   47;  Gaps   11;

Qy        106 YLTS---------SLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNL 154

Db         14 YYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLLNL 72

Qy        155 AIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAI 214

Db         73 AISDLLFLLTVPFWA--HYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAI 130

Qy        215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCH--DVHNTCE 272

Db        131 VHAVFALKARTVTFGVVTSVITWVVAVFASLPRII----------FTTSHRERLHYTCS 179

Qy        273 SSSPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKA 321

Db        180 SHFPYSQYQFWKNFHTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRL 236

Qy        322 SLLILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFL 376

Db        237 IFTIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPII 296

Qy        377 Y-FLMSKTRNHSTAYLTK 393

Db        297 YAFVGEKFRNYLLVFFQK 314


RESULT 28
Q9Z282
ID   Q9Z282      PRELIMINARY;      PRT;   382 AA.
AC   Q9Z282;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor G2A.
GN   G2A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98445372; PubMed=9770487;
RA   Weng Z., Fluckiger A.C., Nisitani S., Wahl M.I., Le L.Q., Hunter C.A.,
RA   Fernal A.A., Le Beau M.M., Witte O.N.;
RT   "A DNA damage and stress inducible G protein-coupled receptor blocks
RT   cells in G2/M.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:12334-12339(1998).
DR   EMBL; AF083442; AAC67542.1; -.
DR   MGD; MGI:1890220; G2a.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01563; G2ARECEPTOR.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   382 AA;  42652 MW;  D4F8CE0370CCD610 CRC64;

  Query Match             15.4%;  Score 331.5;  DB 11;  Length 382;
  Best Local Similarity   27.2%;  Pred. No. 1.4e-20;
  Matches   85;  Conservative   65;  Mismatches  115;  Indels   47;  Gaps   10;

Qy        101 NATMGYLTSSLSTKLIPA----------------IYLLVFVVGVPANAVTLWMLFFRT- 142

Db         11 NTTLG-VTSVLQSTSVPSSETCHVSYEESRVVLVVVYSAVCLLGLPANCLTAWLTLLQVL 69

Qy        143 -RSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSIL 201

Db         70 QRNV-LAVYLFCLSLCELLYISTVPLWIIYIQNQHKWNLGPQACKVTAYIFFCNIYISIL 128

Qy        202 LLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILK-QEYYLVQPD 260

Db        129 LLCCISCDRYMAVVYALESRGHRHQRTAVTISACVILLVGLVNYPVFDMKVEKSFCFEP- 187

Qy        261 ITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDH----RWL 316

Db        188 ----------LRMNSKIAGYHYLRFT-FGFAIPLGILAFTNHQIFRSIKLSDSLSAAQKN 236

Qy        317 WYVKASLLILVIFTICFAPSNIILIIHHA--NYYYNNTDG-------LYFIYLIALCLGS 367

Db        237 KVKRSAIAVVTIFLVCFAPYHVVLLVKAASFSFYQGDMDAVCAFESRLYTVSMVFLCLST 296
```

```
Qy        368 LNSCLDPFLYFL 379
              :||  || :| |
Db        297 VNSVADPIIYVL 308


RESULT 29
Q8TBK4
ID   Q8TBK4      PRELIMINARY;      PRT;   359 AA.
AC   Q8TBK4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Angiotensin receptor 1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC022447; AAH22447.1; -.
KW   Receptor.
SQ   SEQUENCE   359 AA;  41087 MW;  AF90BE7F21E911BA CRC64;

  Query Match             15.3%;  Score 331;  DB 4;  Length 359;
  Best Local Similarity   29.0%;  Pred. No. 1.4e-20;
  Matches   84;  Conservative   51;  Mismatches  121;  Indels   34;  Gaps    6;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTR-SICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              :|| :| ::|||||:  |:: : :::|  :     :||  |||:||  | :|||    |
Db         30 MIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLTLPLWAVYTA 89

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
                | ||  ||:  :     |:| |: || |:||:|||||||   |       | |||
Db         90 MEYRWPFGNYLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTC 149

Qy        234 GLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIP 293
               ::|    |  ||  | :  :::   :|| |   : :  |: |  |   ::    ||| |
Db        150 IIIWLLAGLASLPAIIHRNVFFIENTNITVCAFHYESQNSTLPIGL--GLTKNILGFLFP 207

Qy        294 FVLIIYCYAAIIRTL-NAYD-----HRWLWYVKASLLILVIFTICFAPSNIILIIHHANY 347
              |::|:  |  | : :  ||:      |     |  : |:: |   : |  |   :
Db        208 FLIILTSYTLIWKAVKKAYEIQKNKPRNDDIFKIIMAIVLFFFFSWIPHQIFTFL----- 262

Qy        348 YYNNTDGLYFIYL--------------IALCLGSLNSCLDPFLYFLMSK 382
                | |  : :                | :|:   |:||:|  |  : |
Db        263 -----DVLIQLGIIRDCRIADIVDTAMPITICIAYFNNCLNPLFYGFLGK 307


RESULT 30
Q9TUR6
ID   Q9TUR6      PRELIMINARY;      PRT;   339 AA.
AC   Q9TUR6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   C-C chemokine receptor 5 (Fragment).
GN   CCR5.
OS   Cercopithecus aethiops vervet.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=100936;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kunstman K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA   Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinksy S.;
RT   "Sequences of the CCR5 genes from diverse simian and prosimian
RT   species.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF162026; AAD47781.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  39178 MW;  9DF2A6F446C55AED CRC64;

  Query Match             15.3%;  Score 330.5;  DB 6;  Length 339;
  Best Local Similarity   25.8%;  Pred. No. 1.5e-20;
  Matches   82;  Conservative   69;  Mismatches  124;  Indels   43;  Gaps   10;

Qy        104 MGYLTS---------SLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYT 152
              : | ||          :: :|:| :| |||: |   | : : :|    | :|: | ::
Db          5 INYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLL 63

Qy        153 NLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYL 212
              ||||:| || :|:||    |     | ||  :|:  | :::   :  |  :  ::|:|||
Db         64 NLAISDLLFLLTVPFWA--HYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYL 121

Qy        213 AIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCE 272
              ||||           |: :||  : |        ||  |  :         :  :| ||
Db        122 AIVHAVFALKARTVTFGVVTSVITWVVAVFASLPRIIFTR---------SQREGLHYTCS 172

Qy        273 SSSPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKA 321
              |  |:  | |      : :   | ::| :::: ||: |::||          ||    |:
Db        173 SHFPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRL 229

Qy        322 SLLILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFL 376
              |:::: : :|| ||:|:::    ::     :::: |      :   || : : |:|| :
Db        230 IFTIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCIDPII 289

Qy        377 Y-FLMSKTRNHSTAYLTK 393
              | |:  | ||:   :  |
Db        290 YAFVGEKFRNYLLVFFQK 307


RESULT 31
O93239
ID   O93239      PRELIMINARY;      PRT;   342 AA.
AC   O93239;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CXC chemokine receptor-2.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Cyprinus.
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99367403; PubMed=10436187;
RA   Fujiki K., Shin D., Nakao M., Yano T.;
RT   "Molecular cloning of carp (Cyprinus carpio) CC chemokine, CXC
RT   chemokine receptors, allograft inflammatory factor-1, and natural
RT   killer cell enhancing factor by use of suppression subtractive
RT   hybridization.";
RL   Immunogenetics 49:909-914(1999).
DR   EMBL; AB010713; BAA31470.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01559; DUFFYANTIGEN.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   342 AA;  38481 MW;  8946E5ED5E534B39 CRC64;

  Query Match             15.3%;  Score 330;  DB 13;  Length 342;
```

```
  Best Local Similarity   27.8%;  Pred. No. 1.7e-20;
  Matches   82;  Conservative   57;  Mismatches  122;  Indels   34;  Gaps    9;

Qy        119 IYLLVFVVGVPANAVTLWML-FFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNN 177
              :|: || : :| | | :::      |   | |:  :|||||  | :|||| |  ::  :
Db         32 VYVSVFCLSLPGNMVVIFVVSCMENRRTSTDVYLMHLAIADLFFALTLPFSAA-DIHAGH 90

Qy        178 WVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVW 237
              |:||  :|:  : :    || : ||||||::||||||    :   :|   :| | |||
Db         91 WMFGTFMCKLISGLQEATFYCCVFLLACISVDRYLAIVKATQFLAQKRHLVGIV-CALVW 149

Qy        238 ATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLI 297
                ||  ||  : :| :         ||| : | ||   :::   |      || :|  ::
Db        150 VCAFLLSLPIIVNRQAFITGNMGDYICHD-NVTAESMDSWRMGLRILHHTLGFFLPLAVM 208

Qy        298 IYCY----AAIIRTLNAYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNT- 352
              ::||      : || |:   :    ::  | ::: | :|: | ||: :|         :|
Db        209 MFCYGFTMCTLCRTRNSQKQK---AMRVILSVVLAFIVCWLPFNILELI--------DTL 257

Qy        353 ------------DGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNHSTAYLTK 393
                         | :   :   :   :  ::| || |:  | ||      | |
Db        258 MRGGKVIETCELRDSIDVALYVTQAMAFAHCAINPILYAFIGKKFRNQLLMSLFK 312


RESULT 32
Q9TUU5
ID   Q9TUU5      PRELIMINARY;      PRT;   339 AA.
AC   Q9TUU5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   C-C chemokine receptor 5 (Fragment).
GN   CCR5.
OS   Macaca fuscata (Japanese macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9542;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kunstman K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA   Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinksy S.;
RT   "Sequences of the CCR5 genes from diverse simian and prosimian
RT   species.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF161957; AAD47713.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  39124 MW;  EA9CD7AA531E599A CRC64;

  Query Match             15.3%;  Score 329.5;  DB 6;  Length 339;
  Best Local Similarity   25.9%;  Pred. No. 1.8e-20;
  Matches   82;  Conservative   68;  Mismatches  123;  Indels   43;  Gaps   10;

Qy        106 YLTS---------SLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNL 154
              | ||          :: :|:| :| |||: |   | : : :|    | :|: | ::  ||
Db          7 YYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLLNL 65

Qy        155 AIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAI 214
              ||:| || :|:||    |     | ||  :|:  | :::   :  |  :  ::|:|||||
Db         66 AISDLLFLLTVPFWA--HYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAI 123

Qy        215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESS 274
              ||          |: :||  : |       ||  |  :         :    :| || |
Db        124 VHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTR---------SQREGLHYTCSSH 174

Qy        275 SPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKASL 323
              |:  | |      : :   | ::| :::: ||: |::||          ||    |:
Db        175 FPYSQYQFWKNFQTLKMVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRLIF 231

Qy        324 LILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFLY- 377
              |:::: : :|| ||:|:::    ::     |::: |      :   ||  : |::| :|
Db        232 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCNSSNRLDQAMQVTETLGMTHCCINPIIYA 291

Qy        378 FLMSKTRNHSTAYLTK 393
              |:  | ||:   :  |
Db        292 FVGEKFRNYLLVFFQK 307


RESULT 33
Q9TUR4
ID   Q9TUR4      PRELIMINARY;      PRT;   339 AA.
AC   Q9TUR4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   C-C chemokine receptor 5 (Fragment).
GN   CCR5.
OS   Mandrillus sphinx (Mandrill) (Papio sphinx).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Mandrillus.
OX   NCBI_TaxID=9561;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kunstman K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA   Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinksy S.;
RT   "Sequences of the CCR5 genes from diverse simian and prosimian
RT   species.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF162032; AAD47787.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  39135 MW;  518626A2DEF1BFBE CRC64;

  Query Match             15.3%;  Score 329.5;  DB 6;  Length 339;
  Best Local Similarity   25.9%;  Pred. No. 1.8e-20;
  Matches   82;  Conservative   69;  Mismatches  122;  Indels   43;  Gaps   10;

Qy        106 YLTSSL---------STKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNL 154
              | || |          : :|:| :| |||: |   | : : :|    | :|: | ::  ||
Db          7 YYTSELCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLLNL 65

Qy        155 AIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAI 214
              ||:| || :|:||    |     | ||  ::|:  | :::   :  |  :  ::|:|||||
Db         66 AISDLLFLLTVPFWA--HYAAAQWDFGNIMCQLLTGLYFIGFFSGIFFIILLTIDRYLAI 123

Qy        215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESS 274
              ||          |: :||  : |       ||  |  :         :    :| || |
Db        124 VHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTR---------SQREGLHYTCSSH 174

Qy        275 SPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKASL 323
              |:  | |      : :   | ::| :::: ||: |::||          ||    |:
Db        175 FPYSQYQFWKNFRTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRLIF 231

Qy        324 LILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFLY- 377
              |:::: : :|| ||:|:::    ::     |::: |      :   ||  : |::| :|
Db        232 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYA 291

Qy        378 FLMSKTRNHSTAYLTK 393
              |:  | ||:   :  |
Db        292 FVGEKFRNYLLVFFQK 307
```

```
RESULT 34
Q95NC3
ID   Q95NC3      PRELIMINARY;      PRT;   352 AA.
AC   Q95NC3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   C-C chemokine receptor 5.
GN   CCR5.
OS   Miopithecus talapoin (Talapoin) (Cercopithecus talapoin).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Miopithecus.
OX   NCBI_TaxID=36231;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang Y., Ryder O.A., Zhang Y.;
RT   "Sequence comparison of the CCR5 gene in primates and primate
RT   phylogeny.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF177886; AAK43369.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   352 AA;  40546 MW;  6464152F3E566AE5 CRC64;

  Query Match             15.3%;  Score 329.5;  DB 6;  Length 352;
  Best Local Similarity   25.8%;  Pred. No. 1.9e-20;
  Matches   82;  Conservative   69;  Mismatches  124;  Indels   43;  Gaps   10;

Qy        104 MGYLTS---------SLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYT 152
              : | ||          :: :|:| :| |||: |   | : : :|    | :|: | ::
Db         12 INYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLL 70

Qy        153 NLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYL 212
              ||||:| || :|:|| |    | || :|| | :::  :  |  :  ::|:|||
Db         71 NLAISDLLFLLTVPFWA--HYAAAQWDFGNTMCRLLTGLYFIGFFSGIFFIILLTIDRYL 128

Qy        213 AIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCE 272
              ||||          |: :||  : |       ||  |  :          :    :| ||
Db        129 AIVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTR---------SQREGLHYTCS 179

Qy        273 SSSPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKA 321
              |  |:  | |      : :   | ::| :::: ||: |::||         ||    |:
Db        180 SHFPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRL 236

Qy        322 SLLILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFL 376
              |:::: : :|| ||:|:::    ::     :::: |     :   ||  : |::| :
Db        237 IFTIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPII 296

Qy        377 Y-FLMSKTRNHSTAYLTK 393
              | |:  | ||:   :  |
Db        297 YAFVGEKFRNYLLVFFQK 314


RESULT 35
Q9TUU8
ID   Q9TUU8      PRELIMINARY;      PRT;   339 AA.
AC   Q9TUU8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   C-C chemokine receptor 5 (Fragment).
GN   CCR5.
OS   Cercopithecus diana (Diana monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=36224;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kunstman K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA   Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinksy S.;
RT   "Sequences of the CCR5 genes from diverse simian and prosimian
RT   species.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF161949; AAD47705.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  39049 MW;  6D1A93F66270F3ED CRC64;

  Query Match             15.2%;  Score 328.5;  DB 6;  Length 339;
  Best Local Similarity   25.9%;  Pred. No. 2.2e-20;
  Matches   82;  Conservative   69;  Mismatches  122;  Indels   43;  Gaps   10;

Qy        106 YLTS---------SLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNL 154
              | ||          :: :|:| :| |||: |   | : : :|    | :|: | ::  ||
Db          7 YYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLLNL 65

Qy        155 AIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAI 214
              ||:| || :|:||    |     | ||  :|:  | :::   :  |  :  ::|:|||||
Db         66 AISDLLFLLTVPFWA--HYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAI 123

Qy        215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESS 274
              ||          |: :||  : |       ||  |  :          :    :| || |
Db        124 VHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTR---------SQREGLHYTCSSH 174

Qy        275 SPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKASL 323
              |:  | |      : :   | ::| :::: ||: |::||         ||    |:
Db        175 FPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRLIF 231

Qy        324 LILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFLY- 377
              |:::: : :|| ||:|:::    ::     :::: |     :|  ||  : |::| :|
Db        232 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVAETLGMTHCCINPIIYA 291

Qy        378 FLMSKTRNHSTAYLTK 393
              |:  | ||:   :  |
Db        292 FVGEKFRNYLLVFFQK 307


RESULT 36
Q9TSQ3
ID   Q9TSQ3      PRELIMINARY;      PRT;   339 AA.
AC   Q9TSQ3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   C-C chemokine receptor 5 (Fragment).
GN   CCR5.
OS   Cercopithecus aethiops (Green monkey) (Grivet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=9534;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kunstman K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA   Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinksy S.;
RT   "Sequences of the CCR5 genes from diverse simian and prosimian
RT   species.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF162007; AAD47762.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
```

```
FT   NON_TER       1      1
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  39219 MW;  BFFE9138E82933D0 CRC64;

  Query Match           15.2%;  Score 328.5;  DB 6;  Length 339;
  Best Local Similarity   25.9%;  Pred. No. 2.2e-20;
  Matches   82;  Conservative   68;  Mismatches  123;  Indels   43;  Gaps   10;

Qy        106 YLTS---------SLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNL 154
              | ||          :: :|:| :| ||||: |   | : : :|    | :|: | ::  ||
Db          7 YYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLLNL 65

Qy        155 AIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAI 214
              ||:| || :|:||    |      | ||  :|:  | :::   :  |  :  ::|:|||||
Db         66 AISDLLLFLLTVPFWA--HYAAARWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAI 123

Qy        215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESS 274
              ||          |: :||  : |        || | :          :  :| || |
Db        124 VHAVFALKARTVTFGVVTSVITWVVAVFASLPRIIFTR---------SQREGLHYTCSSH 174

Qy        275 SPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKASL 323
              |:  | |      : :  | ::| :::: ||: |::||         ||    |:
Db        175 FPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRLIF 231

Qy        324 LILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFLY- 377
              |:::: : :|| ||:|:::    ::      :|:: |      :   ||  : |::|.:|
Db        232 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSNSNRLDQAMQVTDTLGMTHCCINPIIYA 291

Qy        378 FLMSKTRNHSTAYLTK 393
              |:  | ||:   :  |
Db        292 FVGEKFRNYLLVFFQK 307


RESULT 37
Q9N0Z0
ID   Q9N0Z0      PRELIMINARY;      PRT;   343 AA.
AC   Q9N0Z0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   STRL33.
OS   Cercocebus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercocebus.
OX   NCBI_TaxID=9531;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20261727; PubMed=10799581;
RA   Pohlmann S., Lee B., Meister S., Krumbiegel M., Leslie G., Doms R.W.,
RA   Kirchhoff F.;
RT   "Simian immunodeficiency virus utilizes human and sooty mangabey but
RT   not rhesus macaque STRL33 for efficient entry.";
RL   J. Virol. 74:5075-5082(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Poehlmann S., Lee B., Meister S., Krumbiegel M., Leslie G., Doms R.W.,
RA   Kirchhoff F.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF237559; AAF68392.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01568; LYMPHOTACTNR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   343 AA;  39589 MW;  A75B7A0751C13455 CRC64;

  Query Match           15.2%;  Score 328.5;  DB 6;  Length 343;
  Best Local Similarity   30.1%;  Pred. No. 2.3e-20;
  Matches   84;  Conservative   54;  Mismatches  116;  Indels   25;  Gaps   10;

Qy        112 STKLIPAIYLLVFVVGVPANAVTLWM-LFFRTRSICTTVFYTNLAIADFLFCVTLPFKIA 170
              |   :| :||:||| |:  |:: | : :|:       | ||  || :|| :|  ||||
Db         31 SKVFLPCMYLVVFVCGLVGNSLVLVISIFYHKLQSLTDVFLVNLPLADLVFVCTLPFWA- 89

Qy        171 YHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF-TYRGLPKH-TY 228
              :   : .|:||:|:|:    ::  | | |:|:| ||:::|:: :|     |    |  |:
Db         90 -YAGIHEWIFGQVMCKTLLGVYTINFYTSMLILTCITVDRFIVVVKATKAYNQQAKRMTW 148

Qy        229 ALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTC--HDVHNTCESSSPFQLYYFISLA 286
              | | |:|    |  ||  |    : |   |   |  ||      |  |   |   ::|
Db        149 GKVICLLIWVISLLVSLPQIIYGNVFNL---DKLICRYHD-----EEISTVVLATQMTL- 199

Qy        287 FFGFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNIILII 342
              || :| : :| ||: ||:||       || |   |   |:: :| :   | |:: :|
Db        200 --GFFLPLLTMIVCYSVIIKTLLHAGGFQKHRSL---KIIFLVMAVFLLTQTPFNLVKLI 254

Qy        343 HHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMS 381
              :: |      ::  ::   :  | :||:| ||  :|
Db        255 RSTHWEYYAMTSFHYTIIVTEAIAYLRACLNPVLYAFVS 293


RESULT 38
Q9MYJ8
ID   Q9MYJ8      PRELIMINARY;      PRT;   355 AA.
AC   Q9MYJ8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Chemokine receptor.
GN   CCR1.
OS   Callithrix jacchus (Common marmoset).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.
OX   NCBI_TaxID=9483;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20153429; PubMed=10686294;
RA   Liang M., Rosser M.P., Ng H.P., May K., Bauman J.G., Islam I.,
RA   Ghannam A., Kretschmer P.J., Pu H., Dunning L., Snider R.M.,
RA   Morrissey M.M., Hesselgesser J., Perez H.D., Horuk R.;
RT   "Species selectivity of a small molecule antagonist for the CCR1
RT   chemokine receptor.";
RL   Eur. J. Pharmacol. 389:41-49(2000).
DR   EMBL; AF127528; AAF36453.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   355 AA;  40928 MW;  2B01C47E9874A2C1 CRC64;

  Query Match           15.2%;  Score 328;  DB 6;  Length 355;
  Best Local Similarity   28.1%;  Pred. No. 2.6e-20;
  Matches   82;  Conservative   63;  Mismatches  121;  Indels   26;  Gaps   10;

Qy        110 SLSTKLIPAIYLLVFVVGVPANAVTLWMLF-FRTRSICTTVFYTNLAIADFLFCVTLPFK 168
              :  ||:| :| ||||:|:  | : : :|  ::       |:::  ||||:| ||  ||||
Db         31 AFGAKLLPPLYSLVFVIGLVGNILVVVVLVQYKRLKNMTSIYLLNLAISDLLFLFTLPFW 90

Qy        169 IAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTY 228
              |:| |   :||||  :|:  :  :|  :|  |  :  ::|:|||||||          |:
Db         91 ISYQLK-TDWVFGNAMCKVLSGFYYTGLYSEIFFIILLTIDRYLAIVHAVFALRARTVTF 149

Qy        229 ALVTCGLVWATVFLYMLP-FFILKQEYYLVQPDITTCHDVHNTC------ESSSPFQLYY 281
              ::|  ::|    |  || :  | ::      :||      | ||      ||   ::|:
Db        150 GVITSIIIWVLAILASLPGLYFAKTQW-----EIT-----HRTCSLHFPHESRQEWKLFQ 199

Qy        282 FISLAFFGFLIPFVLIIYCYAAIIR-TLNAYDHRWLWYVKASLLILVIFTICFAPSN--- 337
              : |   | ::| :::| ||  ||:  |   : :     |:   :|::|| : : | |
Db        200 ALKLNLLGLVLPLLVMIVCYTGIIKILLRRPNEKKSKAVRLIFVIMIIFFLFWTPYNLTT 259

Qy        338 IILIIHHANYYYNNTDG--LYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNH 386
              :| :    : |    | |     :  :  .: |::| :| |:  : | |
```

```
Db        260 LISVFQDFLFTYGCEQGRQLDLAIQVTEMIAYTHCCVNPVIYAFVGERFRKH 311


RESULT 39
Q9TUU6
ID   Q9TUU6      PRELIMINARY;      PRT;   339 AA.
AC   Q9TUU6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   C-C chemokine receptor 5 (Fragment).
GN   CCR5.
OS   Macaca fuscata (Japanese macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9542;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kunstman K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA   Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinksy S.;
RT   "Sequences of the CCR5 genes from diverse simian and prosimian
RT   species.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF161956; AAD47712.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  39080 MW;  4DCA9C72B093216D CRC64;

  Query Match             15.2%;  Score 327.5;  DB 6;  Length 339;
  Best Local Similarity   25.9%;  Pred. No. 2.7e-20;
  Matches   82;  Conservative   68;  Mismatches  123;  Indels   43;  Gaps   10;

Qy        106 YLTS---------SLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNL 154
              | ||          :: :|:| :| |||: |   | : : :|    | :|: | ::  ||
Db          7 YYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLLNL 65

Qy        155 AIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAI 214
              ||:| || :|:||    |     | ||  :|:  | :::   :  |  :  ::|:|||||
Db         66 AISDLLFLLTVPFWA--HYAAARWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAI 123

Qy        215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESS 274
              ||          |: :||  : |       ||  |  :         :    :| || |
Db        124 VHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTR---------SQRGGLHYTCSSH 174

Qy        275 SPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKASL 323
              |:  | |       : :   | ::| :::: ||: |::||          ||    |:
Db        175 FPYSQYQFWKNFQTLKMVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRLIF 231

Qy        324 LILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFLY- 377
               |:::: : :|| ||:|:::    ::      |::: |      :   ||  : |::| :|
Db        232 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCNSSNRLDQAMQVTETLGMTHCCINPIIYA 291

Qy        378 FLMSKTRNHSTAYLTK 393
              |:  | ||:   : |
Db        292 FVGEKFRNYLLVFFQK 307


RESULT 40
Q9TQX3
ID   Q9TQX3      PRELIMINARY;      PRT;   339 AA.
AC   Q9TQX3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   C-C chemokine receptor 5 (Fragment).
GN   CCR5.
OS   Mandrillus sphinx (Mandrill) (Papio sphinx).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Mandrillus.
OX   NCBI_TaxID=9561;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kunstman K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA   Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinksy S.;
RT   "Sequences of the CCR5 genes from diverse simian and prosimian
RT   species.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF161999; AAD47754.1; -.
DR   EMBL; AF161998; AAD47753.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  39119 MW;  066BFC06CD00F0D8 CRC64;

  Query Match             15.2%;  Score 327.5;  DB 6;  Length 339;
  Best Local Similarity   25.6%;  Pred. No. 2.7e-20;
  Matches   81;  Conservative   70;  Mismatches  122;  Indels   43;  Gaps   10;

Qy        106 YLTS---------SLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNL 154
              | ||          :: :|:| :| |||: |   | : : :|    | :|: | ::  ||
Db          7 YYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNILVVLILINCKRLKSM-TDIYLLNL 65

Qy        155 AIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAI 214
              ||:| || :|:||    |     | ||  ::|:  | :::   :  |  :  ::|:|||||
Db         66 AISDLLFLLTVPFWA--HYAAAQWDFGNIMCQLLTGLYFIGFFSGIFFIILLTIDRYLAI 123

Qy        215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESS 274
              ||          |: :||  : |       ||  |  :         :    :| || |
Db        124 VHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTR---------SQREGLHYTCSSH 174

Qy        275 SPFQLYYF------ISLAFFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYVKASL 323
              |:  | |       : :   | ::| :::: ||: |::||          ||    |:
Db        175 FPYSQYQFWKNFRTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHR---AVRLIF 231

Qy        324 LILVIFTICFAPSNIILIIHHANYYY-----NNTDGLYFIYLIALCLGSLNSCLDPFLY- 377
               |:::: : :|| ||:|:::    ::      |::: |      :   ||  : |::| :|
Db        232 TIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYA 291

Qy        378 FLMSKTRNHSTAYLTK 393
              |:  | ||:   : |
Db        292 FVGEKFRNYLLVFFQK 307


Search completed: June 24, 2003, 12:05:38
Job time : 54.5632 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:53:16 ; Search time 14.5178 Seconds
(without alignments)
1165.628 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIKM.....AYLTKXXNDLREQGQPSQRT 408

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1998	92.6	374	1	PAR3_HUMAN	O00254 homo sapien
2	1425.5	66.1	369	1	PAR3_MOUSE	O08675 mus musculu
3	1414	65.6	368	1	PAR3_RAT	Q920e1 rattus norv
4	606.5	28.1	399	1	PAR2_MOUSE	P55086 mus musculu
5	599.5	27.8	420	1	PAR1_XENLA	P47749 xenopus lae
6	595	27.6	397	1	PAR2_RAT	Q63645 rattus norv
7	589.5	27.3	397	1	PAR2_HUMAN	P55085 homo sapien
8	535	24.8	395	1	PAR4_RAT	Q920e0 rattus norv
9	531.5	24.6	396	1	PAR4_MOUSE	O88634 mus musculu
10	522.5	24.2	430	1	PAR1_MOUSE	P30558 mus musculu
11	521.5	24.2	385	1	PAR4_HUMAN	Q96ri0 homo sapien
12	520.5	24.1	425	1	PAR1_PAPHA	P56488 papio hamad
13	519	24.1	425	1	PAR1_HUMAN	P25116 homo sapien
14	505.5	23.4	432	1	PAR1_RAT	P26824 rattus norv
15	501	23.2	428	1	PAR1_CRILO	Q00991 cricetulus
16	428	19.8	537	1	P2Y8_XENLA	P79928 xenopus lae
17	426	19.7	362	1	P2YR_MELGA	P49652 meleagris g
18	424	19.7	362	1	P2YR_CHICK	P34996 gallus gall
19	410.5	19.0	361	1	EBI2_HUMAN	P32249 homo sapien
20	410.5	19.0	367	1	GP17_HUMAN	Q13304 homo sapien
21	409.5	19.0	373	1	P2YR_RAT	P49651 rattus norv
22	406.5	18.8	308	1	P2Y5_CHICK	P32250 gallus gall
23	399	18.5	373	1	P2YR_HUMAN	P47900 homo sapien
24	398	18.5	373	1	P2YR_BOVIN	P48042 bos taurus
25	397.5	18.4	373	1	P2YR_MOUSE	P49650 mus musculu
26	384.5	17.8	370	1	P2Y9_HUMAN	Q99677 homo sapien
27	382	17.7	365	1	P2Y4_HUMAN	P51582 homo sapien
28	380.5	17.6	309	1	CLT2_RAT	Q924t9 rattus norv
29	379.5	17.6	345	1	CLT2_PIG	Q95n03 sus scrofa
30	378	17.5	361	1	P2Y4_RAT	O35811 rattus norv
31	376.5	17.5	346	1	CLT2_HUMAN	Q9ns75 homo sapien
32	376	17.4	344	1	P2Y5_HUMAN	P43657 homo sapien
33	375.5	17.4	340	1	CLT1_PIG	Q95n02 sus scrofa
34	372.5	17.3	337	1	CLT1_HUMAN	Q9y271 homo sapien
35	371.5	17.2	352	1	CLT1_MOUSE	Q99ja4 mus musculu
36	370	17.2	361	1	P2Y4_MOUSE	Q9jjs7 mus musculu
37	366	17.0	377	1	P2Y2_HUMAN	P41231 homo sapien
38	365.5	16.9	339	1	CLT1_RAT	Q924t8 rattus norv
39	365.5	16.9	374	1	P2Y2_RAT	P41232 rattus norv
40	365.5	16.9	381	1	GP34_HUMAN	Q9upc5 homo sapien
41	364.5	16.9	328	1	P2Y3_CHICK	Q98907 gallus gall
42	363.5	16.9	355	1	IL8A_RABIT	P21109 oryctolagus
43	361.5	16.8	309	1	CLT2_MOUSE	Q920a1 mus musculu
44	361.5	16.8	363	1	AG2S_XENLA	P35373 xenopus lae
45	361	16.7	373	1	P2Y2_MOUSE	P35383 mus musculu

ALIGNMENTS

```
RESULT 1
PAR3_HUMAN
ID   PAR3_HUMAN      STANDARD;      PRT;   374 AA.
AC   O00254;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-
DE   like 2) (Coagulation factor II receptor-like 2).
GN   F2RL2 OR PAR3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF THR-39 AND PHE-40.
RX   MEDLINE=97242411; PubMed=9087410;
RA   Ishihara H., Connolly A.J., Zeng D., Kahn M.L., Zheng Y.-W.,
RA   Timmons C., Tram T., Coughlin S.R.;
RT   "Protease-activated receptor 3 is a second thrombin receptor in
RT   humans.";
RL   Nature 386:502-506(1997).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANTS SER-15; VAL-177 AND ASP-250.
RA   Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=98279023; PubMed=9614115;
RA   Schmidt V.A., Nierman W.C., Maglott D.R., Cupit L.D., Moskowitz K.A.,
RA   Wainer J.A., Bahou W.F.;
RT   "The human proteinase-activated receptor-3 (PAR-3) gene.
RT   Identification within a PAR gene cluster and characterization in
RT   vascular endothelial cells and platelets.";
RL   J. Biol. Chem. 273:15061-15068(1998).
RN   [4]
RP   FUNCTION.
RX   MEDLINE=99178892; PubMed=10079109;
RA   Kahn M.L., Nakanishi-Matsui M., Shapiro M.J., Ishihara H.,
RA   Coughlin S.R.;
RT   "Protease-activated receptors 1 and 4 mediate activation of human
RT   platelets by thrombin.";
RL   J. Clin. Invest. 103:879-887(1999).
CC   -!- FUNCTION: Receptor for activated thrombin coupled to G proteins
CC       that stimulate phosphoinositide hydrolysis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Highest expression in the megakaryocytes of
CC       the bone marrow, lower in mature megakaryocytes, in platelets and
CC       in a variety of other tissues such as heart and gut.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U92971; AAC51218.1; -.
DR   EMBL; AF374726; AAK51564.1; -.
DR   HSSP; P34996; 1DDD.
DR   Genew; HGNC:3539; F2RL2.
DR   MIM; 601919; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation; Polymorphism.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     38       REMOVED FOR RECEPTOR ACTIVATION (BY
FT                                SIMILARITY).
FT   CHAIN        39    374       PROTEINASE ACTIVATED RECEPTOR 3.
FT   DOMAIN       39     94       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     95    120       1 (POTENTIAL).
FT   DOMAIN      121    128       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    129    148       2 (POTENTIAL).
FT   DOMAIN      149    167       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    168    189       3 (POTENTIAL).
FT   DOMAIN      190    206       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    207    230       4 (POTENTIAL).
FT   DOMAIN      231    259       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    261    280       5 (POTENTIAL).
FT   DOMAIN      281    297       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    298    322       6 (POTENTIAL).
FT   DOMAIN      323    336       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    337    361       7 (POTENTIAL).
FT   DOMAIN      362    374       CYTOPLASMIC (POTENTIAL).
FT   SITE         38     39       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DISULFID    166    245       BY SIMILARITY.
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    331    331       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      15     15       L -> S.
FT                                /FTId=VAR_012849.
FT   VARIANT     177    177       M -> V.
FT                                /FTId=VAR_012850.
FT   VARIANT     250    250       N -> D.
FT                                /FTId=VAR_012851.
FT   MUTAGEN      39     39       T->P: NO PROTEOLYTIC CLEAVAGE (BY
FT                                THROMBIN).
FT   MUTAGEN      40     40       F->A: ALTERED SIGNAL UPON THROMBIN
FT                                CLEAVAGE.
SQ   SEQUENCE   374 AA;  42508 MW;  C45C15A695DD1ABB CRC64;
```

Query Match 92.6%; Score 1998; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. No. 4.4e-102;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          20 MKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPNSFEEFPPFSALEGWT 79
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPNSFEEFPPFSALEGWT 60

Qy          80 GATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLF 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLF 120

Qy         140 FRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCS 199
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 FRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCS 180

Qy         200 ILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQP 259
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQP 240

Qy         260 DITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWYV 319
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 DITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWYV 300

Qy         320 KASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFL 379
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 KASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFL 360

Qy         380 MSKTRNHSTAYLTK 393
               ||||||||||||||
Db         361 MSKTRNHSTAYLTK 374
```

```
RESULT 2
PAR3_MOUSE
ID   PAR3_MOUSE     STANDARD;      PRT;   369 AA.
AC   O08675;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-
DE   like 2) (Coagulation factor II receptor-like 2).
GN   F2RL2 OR PAR3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=97242411; PubMed=9087410;
RA   Ishihara H., Connolly A.J., Zeng D., Kahn M.L., Zheng Y.-W.,
RA   Timmons C., Tram T., Coughlin S.R.;
RT   "Protease-activated receptor 3 is a second thrombin receptor in
RT   humans.";
RL   Nature 386:502-506(1997).
CC   -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis. May play a
CC       role in platelets activation.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U92972; AAC53137.1; -.
DR   HSSP; P34996; 1DDD.
DR   MGD; MGI:1298208; F2rl2.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation.
FT   SIGNAL        1     19       POTENTIAL.
FT   PROPEP       20     37       REMOVED FOR RECEPTOR ACTIVATION (BY
FT                                SIMILARITY).
FT   CHAIN        38    369       PROTEINASE ACTIVATED RECEPTOR 3.
FT   DOMAIN       38     93       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     94    119       1 (POTENTIAL).
FT   DOMAIN      120    127       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    128    147       2 (POTENTIAL).
FT   DOMAIN      148    166       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    167    188       3 (POTENTIAL).
```

```
FT   DOMAIN      189    205       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    206    229       4 (POTENTIAL).
FT   DOMAIN      230    258       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    260    279       5 (POTENTIAL).
FT   DOMAIN      280    296       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    297    321       6 (POTENTIAL).
FT   DOMAIN      322    335       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    336    360       7 (POTENTIAL).
FT   DOMAIN      361    369       CYTOPLASMIC (POTENTIAL).
FT   SITE         37     38       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DISULFID    165    244       BY SIMILARITY.
FT   CARBOHYD     24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     81     81       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   369 AA;  41707 MW;  F4166BF766D073DB CRC64;

  Query Match             66.1%;  Score 1425.5;  DB 1;  Length 369;
  Best Local Similarity   71.6%;  Pred. No. 4.4e-71;
  Matches  260;  Conservative   44;  Mismatches   58;  Indels    1;  Gaps    1;

Qy         20 MKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEGWT 79
              || ||  ||||| |  | ||||: | ::| ||||| ||:| | | |:|||| | :||||
Db          1 MKILILVAAGLLFLPVTVCQSGI-NVSDNSAKPTLTIKSFNGGPQNTFEEFPLSDIEGWT 59

Qy         80 GATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLF 139
              ||| |:| :|||:| | ||| |||:||| |||||::|||||:|:||||||:| |||| |
Db         60 GATTTIKAECPEDSISTLHVNNATIGYLRSSLSTQVIPAIYILLFVVGVPSNIVTLWKLS 119

Qy        140 FRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCS 199
               ||:||   :|:||||||| |||||||||||||||||||||:|| |||:|||||||:
Db        120 LRTKSISLVIFHTNLAIADLLFCVTLPFKIAYHLNGNNWVFGEVMCRITTVVFYGNMYCA 179

Qy        200 ILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQP 259
              ||:| |: ||||||  |||||: ||| :::|: ||:||  |||||||| ||||||:||
Db        180 ILILTCMGINRYLATAHPFTYQKLPKRSFSLLMCGIVWVMVFLYMLPFVILKQEYHLVHS 239

Qy        260 DITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWYV 319
              :||||||| : ||| | |: |||:||||||||||||:||:||  :|   | : |   || |:
Db        240 EITTCHDVVDACESPSSFRFYYFVSLAFFGFLIPFVIIIFCYTTLIHKLKSKDRIWLGYI 299

Qy        320 KASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFL 379
              || |||||||||||:||||:||||||||:||| |||:|||||||||||||||||||||:
Db        300 KAVLLILVIFTICFAPTNIILVIHHANYYYHNTDSLYFMYLIALCLGSLNSCLDPFLYFV 359

Qy        380 MSK 382
              |||
Db        360 MSK 362


RESULT 3
PAR3_RAT
ID   PAR3_RAT       STANDARD;      PRT;   368 AA.
AC   Q920E1;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-
DE   like 2) (Coagulation factor II receptor-like 2).
GN   F2RL2 OR PAR3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA   Chien E.K., Marietti S., Mendoza J., Phillippe M.;
RT   "Cloning of the rat protease activated receptor isoforms 3 and 4.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for activated thrombin coupled to G proteins
CC       that stimulate phosphoinositide hydrolysis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF310076; AAL26789.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     37       REMOVED FOR RECEPTOR ACTIVATION (BY
FT                                SIMILARITY).
FT   CHAIN        38    368       PROTEINASE ACTIVATED RECEPTOR 3.
FT   DOMAIN       38     93       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     94    114       1 (POTENTIAL).
FT   DOMAIN      115    123       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    124    144       2 (POTENTIAL).
FT   DOMAIN      145    166       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    167    187       3 (POTENTIAL).
FT   DOMAIN      188    208       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    209    229       4 (POTENTIAL).
FT   DOMAIN      230    257       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    258    278       5 (POTENTIAL).
FT   DOMAIN      279    300       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    301    321       6 (POTENTIAL).
FT   DOMAIN      322    338       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    339    359       7 (POTENTIAL).
FT   DOMAIN      360    368       CYTOPLASMIC (POTENTIAL).
FT   SITE         37     38       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DISULFID    164    243       BY SIMILARITY.
FT   CARBOHYD     80     80       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   368 AA;  41795 MW;  CE0E94EDA3B80EF1 CRC64;

  Query Match             65.6%;  Score 1414;  DB 1;  Length 368;
  Best Local Similarity   70.4%;  Pred. No. 1.9e-70;
  Matches  262;  Conservative   41;  Mismatches   65;  Indels    4;  Gaps    2;

Qy         18 IKMKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEG 77
              ::|| ||     || |  | |||||:: ::| |   |  ::| |   :|||||| | :||
Db          1 MEMKVLILVGVRLLFLPTTVCQSGMKHVSDNSA---LTAESFNG-NEHSFEEFPLSDIEG 56

Qy         78 WTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWM 137
              ||||| |:| |||||| : ||| ||||||| ||||||:|||||:||||:||||| ||||
Db         57 WTGATTTIKAKCPEESITTLHVNNATMGYLRSSLSTKVIPAIYILVFVIGVPANIVTLWK 116

Qy        138 LFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMY 197
              |  ||:|||  :|:||||||| ||||||||||||||||:||||||:|| ||| ||||||
Db        117 LSSRTKSICLVIFHTNLAIADLLFCVTLPFKIAYHLNGNDWVFGEVMCRVTTVAFYGNMY 176

Qy        198 CSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLV 257
              |:||:| |: |||||| |||||| |||  : |: ||:||  | |||||  ||||||:||
Db        177 CAILILTCMGINRYLATVHPFTYRKLPKRNFTLLMCGVVWVMVVLYMLPLAILKQEYHLV 236

Qy        258 QPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLW 317
              || ||||||||:|||| ||| |||:||||||||||||: ::||  :|  ||| | :||
Db        237 QPGITTCHDVHDTCESPLPFQFYYFVSLAFFGFLIPFVVSVFCYTTLIHKLNAQDRKWLR 296

Qy        318 YVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLY 377
              |:|| |||||||||||||:|||||||||||||:||| |||:||||||||||||||||||
Db        297 YIKAVLLILVIFTICFAPTNIILIIHHANYYYSNTDSLYFMYLIALCLGSLNSCLDPFLY 356

Qy        378 FLMSKTRNHSTA 389
              |:|||  :  |:
Db        357 FIMSKIVDQLTS 368
```

```
RESULT 4
PAR2_MOUSE
ID   PAR2_MOUSE     STANDARD;      PRT;   399 AA.
AC   P55086;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-
DE   like 1) (Coagulation factor II receptor-like 1).
GN   F2RL1 OR PAR2 OR GPR11 OR GPCR11.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95197620; PubMed=7890726;
RA   Nystedt S., Larsson A.-K., Aaberg H., Sundelin J.;
RT   "The mouse proteinase-activated receptor-2 cDNA and gene. Molecular
RT   cloning and functional expression.";
RL   J. Biol. Chem. 270:5950-5955(1995).
CC   -!- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z48043; CAA88097.1; -.
DR   HSSP; P34996; 1DDD.
DR   MGD; MGI:101910; F2rl1.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL       1     25       POTENTIAL.
FT   PROPEP      26     38       REMOVED FOR RECEPTOR ACTIVATION
FT                               (BY SIMILARITY).
FT   CHAIN       39    399       PROTEINASE ACTIVATED RECEPTOR 2.
FT   DOMAIN      39     77       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    78    103       1 (POTENTIAL).
FT   DOMAIN     104    112       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   113    132       2 (POTENTIAL).
FT   DOMAIN     133    151       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   152    173       3 (POTENTIAL).
FT   DOMAIN     174    192       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   193    213       4 (POTENTIAL).
FT   DOMAIN     214    243       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   244    262       5 (POTENTIAL).
FT   DOMAIN     263    287       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   288    310       6 (POTENTIAL).
FT   DOMAIN     311    325       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   326    349       7 (POTENTIAL).
FT   DOMAIN     350    399       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     385    394       POLY-SER.
FT   SITE        38     39       CLEAVAGE (BY TRYPSIN) (POTENTIAL).
FT   DISULFID   150    228       BY SIMILARITY.
FT   CARBOHYD    33     33       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   224    224       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   399 AA;  44752 MW;  A93749425ED0B194 CRC64;

  Query Match             28.1%;  Score 606.5;  DB 1;  Length 399;
  Best Local Similarity   35.4%;  Pred. No. 1.1e-26;
  Matches  137;  Conservative   69;  Mismatches  126;  Indels   55;  Gaps   12;

Qy         22 ALIFAAAGLLLLLPTFCQSGMEN--DTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEGWT 79
              :| :   |: || :   |  ||     | :|    |      ||             |
Db          5 SLAWLLGGITLLAASVSCSRTENLAPGRNNSKGRSLIGRLETQPP------------IT 51

Qy         80 GATITVK--IKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWM 137
              |  : |:      | |||           ||  |:|  :| :|::|||:|:|:| : ||:
Db         52 GKGVPVEPGFSIDEFSAS----------ILTGKLTTVFLPVVYIIVFVIGLPSNGMALWI 101

Qy        138 LFFRTRSICTTVFY-TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNM 196
               |||:     | | |||:|| |  :  | ||:|||:|||||:|| ||:     |||||
Db        102 FLFRTKKKHPAVIYMANLALADLLSVIWFPLKISYHLHGNNWVYGEALCKVLIGFFYGNM 161

Qy        197 YCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKQE 253
              ||||| : |:|: ||  ||:|    | |:   | :  |:   :|  :||  :| :::||
Db        162 YCSILFMTCLSVQRYWVIVNPM---GHPRKK-ANIAVGVSLAIWLLIFLVTIPLYVMKQT 217

Qy        254 YYLVQPDITTCHDVHNTCESSSPFQLY------YFISLAFFGFLIPFVLIIYCYAAIIRT 307
               |:   :|||||||       | ::       ||:|||   || | :|    |  :|:|
Db        218 IYIPALNITTCHDV-------LPEEVLVGDMFNYFLSLAIGVFLFPALLTASAYVLMIKT 270

Qy        308 LNA------YDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLI 361
              | :       : :    ::  : :| :: |||||||::|::|:          :| :||:
Db        271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNLLLVVHYFLIKTQRQSHVYALYLV 330

Qy        362 ALCLGSLNSCLDPFLYFLMSKT-RNHS 387
              |||| :||||:|||:|: :||  |:|:
Db        331 ALCLSTLNSCIDPFVYYFVSKDFRDHA 357


RESULT 5
PAR1_XENLA
ID   PAR1_XENLA     STANDARD;      PRT;   420 AA.
AC   P47749;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94195429; PubMed=8145852;
RA   Gerszten R.E., Chen J., Ishii M., Ishii K., Nanevicz T.,
RA   Turck C.W., Vu T.-K.H., Coughlin S.R.;
RT   "Specificity of the thrombin receptor for agonist peptide is defined
RT   by its extracellular surface.";
RL   Nature 368:648-651(1994).
CC   -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U09632; AAA18498.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
```

```
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation.
FT   SIGNAL        1     20       POTENTIAL.
FT   PROPEP       21     42       REMOVED FOR RECEPTOR ACTIVATION.
FT   CHAIN        43    420       PROTEINASE ACTIVATED RECEPTOR 1.
FT   DOMAIN       43    101       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    102    127       1 (POTENTIAL).
FT   DOMAIN      128    136       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    137    156       2 (POTENTIAL).
FT   DOMAIN      157    175       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    176    197       3 (POTENTIAL).
FT   DOMAIN      198    217       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    218    238       4 (POTENTIAL).
FT   DOMAIN      239    267       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    268    287       5 (POTENTIAL).
FT   DOMAIN      288    310       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    311    333       6 (POTENTIAL).
FT   DOMAIN      334    345       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    346    369       7 (POTENTIAL).
FT   DOMAIN      370    420       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      295    300       POLY-SER.
FT   SITE         42     43       CLEAVAGE (BY THROMBIN).
FT   DISULFID    174    253       BY SIMILARITY.
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     86     86       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   MUTAGEN      42     42       K->A: LACK OF ACTIVATION BY THROMBIN.
SQ   SEQUENCE   420 AA;  47435 MW;  D5163F56AFE12372 CRC64;

  Query Match             27.8%;  Score 599.5;  DB 1;  Length 420;
  Best Local Similarity   34.6%;  Pred. No. 2.7e-26;
  Matches  128;  Conservative   68;  Mismatches  149;  Indels   25;  Gaps    8;

Qy         30 LLLLLPT-------FCQSGMENDTNNLAKPTLPIKTFR--GAPPNSFEEFPFSAL-EGWT 79
              ||||| |       | :  :         : |||||      : ||| |:  | |
Db          8 LLLLLLTLLGAMGSLCLANSDTQAKGAHSNNMTIKTFRIFDDSESEFEEIPWDELDESGE 67

Qy         80 GATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLF 139
              |:    :          ::      ||:|   || :|::| :||:||:| | : : :
Db         68 GSGDQAPVSRSARKPIRRNITKEAEQYLSSQWLTKFVPSLYTVVFIVGLPLNLLAIIIFL 127

Qy        140 FRTRSICTTVFY-TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYC 198
              |: :     | |  ||||||   |   ||||||||:||:|:||  :||  | ||| ||||
Db        128 FKMKVRKPAVVYMLNLAIADVFFVSVLPFKIAYHLSGNDWLFGPGMCRIVTAIFYCNMYC 187

Qy        199 SILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQ 258
              |:||:| ||::|:||:|:|         : | : |  :|       :|  : :|   : :
Db        188 SVLLIASISVDRFLAVVYPMHSLSWRTMSRAYMACSFIWLISIASTIPLLVTEQTQKIPR 247

Qy        259 PDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYD------ 312
              ||||||||  :  :    | :||| |      | :||::   ||  |||:|::
Db        248 LDITTCHDVLD-LKDLKDFYIYYFSSFCLLFFFVPFIITTICYIGIIRSLSSSSIENSCK 306

Qy        313 -HRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSC 371
               | |:    ::::| :| ||| |:|:: : |   |     : ||| |::: |:||:: |
Db        307 KTRALF---LAVVVLCVFIICFGPTNVLFLTH---YLQEANEFLYFAYILSACVGSVSCC 360

Qy        372 LDPFLYFLMS 381
              ||| :|:  |
Db        361 LDPLIYYYAS 370


RESULT 6
PAR2_RAT
ID   PAR2_RAT       STANDARD;      PRT;   397 AA.
AC   Q63645;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-
DE   like 1) (Coagulation factor II receptor-like 1).
GN   F2RL1 OR PAR2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Intestine, and Kidney;
RX   MEDLINE=96358009; PubMed=8762073;
RA   Saifeddine M., Al-Ani B., Cheng C.H., Wang L., Hollenberg M.D.;
RT   "Rat proteinase-activated receptor-2 (PAR-2): cDNA sequence and
RT   activity of receptor-derived peptides in gastric and vascular
RT   tissue.";
RL   Br. J. Pharmacol. 118:521-530(1996).
CC   -!- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U61373; AAC52703.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   PROPEP       26     36       REMOVED FOR RECEPTOR ACTIVATION
FT                                (BY SIMILARITY).
FT   CHAIN        37    397       PROTEINASE ACTIVATED RECEPTOR 2.
FT   DOMAIN       37     75       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     76    101       1 (POTENTIAL).
FT   DOMAIN      102    110       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    111    130       2 (POTENTIAL).
FT   DOMAIN      131    149       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    150    171       3 (POTENTIAL).
FT   DOMAIN      172    190       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    191    211       4 (POTENTIAL).
FT   DOMAIN      212    241       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    242    260       5 (POTENTIAL).
FT   DOMAIN      261    285       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    286    308       6 (POTENTIAL).
FT   DOMAIN      309    323       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    324    347       7 (POTENTIAL).
FT   DOMAIN      348    397       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      383    392       POLY-SER.
FT   SITE         36     37       CLEAVAGE (BY TRYPSIN) (POTENTIAL).
FT   DISULFID    148    226       BY SIMILARITY.
FT   CARBOHYD     23     23       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    222    222       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   397 AA;  44440 MW;  27DE0C3ACABC798D CRC64;

  Query Match             27.6%;  Score 595;  DB 1;  Length 397;
  Best Local Similarity   33.4%;  Pred. No. 4.6e-26;
  Matches  134;  Conservative   71;  Mismatches  128;  Indels   68;  Gaps   12;

Qy         22 ALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRG---APPNSFEEFPFSALEGW 78
              :| :   |: ||      |   | | |   |    ::  |    ||      | :
Db          5 SLAWLLGGITLL----AASASCNRTVNAPGPNSKGRSLIGRLDTPP------PITG---- 50

Qy         79 TGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWML 138
              || :       | |||           ||  |:|  :| ||::|||:|:|:| : ||:
Db         51 KGAPVEPGFSVDEFSAS----------VLTGKLTTVFLPVIYIIVFVIGLPSNGMALWVF 100
```

```
Qy        139 FFRTRSICTTVFY-TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMY 197
              ||||:     |:|  |||:|| |  :  | ||:|||:||:| :|: ||: .    ||||||
Db        101 FFRTKKKHPAVIYMANLALADLLSVIWFPLKISYHLHGNDWTYGDALCKVLIGFFYGNMY 160

Qy        198 CSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKQEY 254
              |||| : |:|: ||  ||:|  :        | :  |:   :|  :||  :| ::::|
Db        161 CSILFMTCLSVQRYWVIVNPMGH----SRKRANIAVGVSLAIWLLIFLVTIPLYVMRQTI 216

Qy        255 YLVQPDITTCHDVHNTCESSSPFQLY------YFISLAFFGFLIPFVLIIYCYAAIIRTL 308
              |:   :|||||||        | ::       ||:|||   || | :|    |  :|:||
Db        217 YIPALNITTCHDV-------LPEEVLVGDMFSYFLSLAIGVFLFPALLTASAYVLMIKTL 269

Qy        309 NA------YDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIA 362
              :       : :    ::  : :| :: |||||||::|::|:          :| :||:|
Db        270 RSSAMDEHSEKKRRRAIRLIITVLSMYFICFAPSNVLLVVHYFLIKSQRQSHVYALYLVA 329

Qy        363 LCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKKXNDLREQGQ 403
              ||| :||||:|||:|: :||              | |:| :
Db        330 LCLSTLNSCIDPFVYYFVSK-------------DFRDQAR 356
```

```
RESULT 7
PAR2_HUMAN
ID   PAR2_HUMAN     STANDARD;      PRT;   397 AA.
AC   P55085; Q13317; Q13346;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-
DE   like 1) (Coagulation factor II receptor-like 1).
GN   F2RL1 OR PAR2 OR GPR11.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96048032; PubMed=7556175;
RA   Nystedt S., Emilsson K., Larsson A.-K., Stroembeck B., Sundelin J.;
RT   "Molecular cloning and functional expression of the gene encoding the
RT   human proteinase-activated receptor 2.";
RL   Eur. J. Biochem. 232:84-89(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=96177879; PubMed=8615752;
RA   Boehm S.K., Kong W., Broemme D., Smeekens S.P., Anderson D.C.,
RA   Connolly A.J., Kahn M.L., Nelken N.A., Coughlin S.R., Payan D.G.,
RA   Bunnett N.W.;
RT   "Molecular cloning, expression and potential functions of the human
RT   proteinase-activated receptor-2.";
RL   Biochem. J. 314:1009-1016(1996).
RN   [3]
RP   SEQUENCE OF 29-397 FROM N.A.
RX   MEDLINE=96379236; PubMed=8784787;
RA   Kahn M.L., Ishii K., Kuo W.L., Piper M., Connolly A.J., Shi Y.P.,
RA   Wu R., Lin C.C., Coughlin S.R.;
RT   "Conserved structure and adjacent location of the thrombin receptor
RT   and protease-activated receptor 2 genes define a protease-activated
RT   receptor gene cluster.";
RL   Mol. Med. 2:349-357(1996).
RN   [4]
RP   SEQUENCE FROM N.A., AND VARIANTS PHE-21; GLN-270 AND ALA-291.
RA   Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary, and Pancreas;
RA   Strausberg R.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis. May have a
CC       role in the regulation of vascular tone.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN TISSUES WITH ESPECIALLY
CC       HIGH LEVELS IN PANCREAS, LIVER, KIDNEY, SMALL INTESTINE, AND
CC       COLON. MODERATE EXPRESSION IS DETECTED IN MANY ORGANS, BUT NONE IN
CC       BRAIN OR SKELETAL MUSCLE.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z49993; CAA90290.1; -.
DR   EMBL; Z49994; CAA90290.1; JOINED.
DR   EMBL; U34038; AAB47871.1; -.
DR   EMBL; U36753; AAA90957.1; -.
DR   EMBL; AF400075; AAK77914.1; -.
DR   EMBL; BC012453; AAH12453.1; -.
DR   EMBL; BC018130; AAH18130.1; -.
DR   HSSP; P34996; 1DDD.
DR   Genew; HGNC:3538; F2RL1.
DR   MIM; 600933; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Polymorphism.
FT   SIGNAL        1     25       POTENTIAL.
FT   PROPEP       26     36       REMOVED FOR RECEPTOR ACTIVATION
FT                                (BY SIMILARITY).
FT   CHAIN        37    397       PROTEINASE ACTIVATED RECEPTOR 2.
FT   DOMAIN       37     75       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     76    101       1 (POTENTIAL).
FT   DOMAIN      102    110       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    111    130       2 (POTENTIAL).
FT   DOMAIN      131    149       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    150    171       3 (POTENTIAL).
FT   DOMAIN      172    190       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    191    211       4 (POTENTIAL).
FT   DOMAIN      212    241       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    242    260       5 (POTENTIAL).
FT   DOMAIN      261    285       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    286    308       6 (POTENTIAL).
FT   DOMAIN      309    323       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    324    347       7 (POTENTIAL).
FT   DOMAIN      348    397       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      383    390       POLY-SER.
FT   SITE         36     37       CLEAVAGE (BY TRYPSIN) (POTENTIAL).
FT   DISULFID    148    226       BY SIMILARITY.
FT   CARBOHYD     30     30       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    222    222       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      21     21       S -> F.
FT                                /FTId=VAR_012846.
FT   VARIANT     270    270       R -> Q.
FT                                /FTId=VAR_012847.
FT   VARIANT     291    291       T -> A.
FT                                /FTId=VAR_012848.
FT   CONFLICT    138    138       G -> A (IN REF. 2).
FT   CONFLICT    291    291       T -> S (IN REF. 5; AAH18130).
SQ   SEQUENCE   397 AA;  44126 MW;  F1A4E1D5AB9B362B CRC64;

  Query Match           27.3%;  Score 589.5;  DB 1;  Length 397;
  Best Local Similarity   35.0%;  Pred. No. 9.1e-26;
  Matches  137;  Conservative   65;  Mismatches  138;  Indels   51;  Gaps   13;
```

```
Qy         28 AGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEGWTGATITVKI 87
Db         12 AAILLAASLSC-SGTIQGTNRSSKGRSLIGKVDGT----------SHVTG-KGVTVETVF 59

Qy         88 KCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICT 147
Db         60 SVDEFSAS----------VLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHP 109

Qy        148 TVFY-TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACI 206
Db        110 AVIYMANLALADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCL 169

Qy        207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHD 266
Db        170 SVQRYWVIVNPMGH-SRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHD 228

Qy        267 VHNTCESSSPFQLY------YFISLAFFGFLIPFVLIIYCYAAIIRTLNA------YDHR 314
Db        229 V-------LPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKK 281

Qy        315 WLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDG---LYFIYLIALCLGSLNSC 371
Db        282 RKRAIKLIVTVLAMYLICFTPSNLLLVVH---YFLIKSQGQSHVYALYIVALCLSTLNSC 338

Qy        372 LDPFLYFLMSKT-RNHS-TAYLTKXXNDLRE 400
Db        339 IDPFVYYFVSHDFRDHAKNALLCRSVRTVKQ 369
```

```
RESULT 8
PAR4_RAT
ID   PAR4_RAT       STANDARD;      PRT;   395 AA.
AC   Q920E0;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-
DE   like 3) (Coagulation factor II receptor-like 3).
GN   F2RL3 OR PAR4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Duodenum;
RA   Hoogerwerf W.A., Lee-Hellmich H., Pasricha P.J.;
RT   "Proteinase activated receptor 4 from rat duodenal library.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Myometrium;
RA   Chien E.K., Marietti S., Mendoza J., Phillippe M.;
RT   "Cloning of the rat protease activated receptor isoforms 3 and 4.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for activated thrombin or trypsin coupled to G
CC       proteins that stimulate phosphoinositide hydrolysis. May play a
CC       role in platelets activation.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF310216; AAL26790.1; -.
DR   EMBL; AF269246; AAK58604.2; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation.
FT   SIGNAL        1     16       POTENTIAL.
FT   PROPEP       17     58       REMOVED FOR RECEPTOR ACTIVATION
FT                                (BY SIMILARITY).
FT   CHAIN        59    395       PROTEINASE ACTIVATED RECEPTOR 4.
FT   DOMAIN       59     93       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     94    114       1 (POTENTIAL).
FT   DOMAIN      115    119       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    120    140       2 (POTENTIAL).
FT   DOMAIN      141    161       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    162    182       3 (POTENTIAL).
FT   DOMAIN      183    203       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    204    224       4 (POTENTIAL).
FT   DOMAIN      225    254       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    255    275       5 (POTENTIAL).
FT   DOMAIN      276    295       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    296    316       6 (POTENTIAL).
FT   DOMAIN      317    330       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    331    354       7 (POTENTIAL).
FT   DOMAIN      355    395       CYTOPLASMIC (POTENTIAL).
FT   SITE         58     59       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN       98    101       POLY-VAL.
FT   DOMAIN      373    376       POLY-SER.
FT   DISULFID    160    239       BY SIMILARITY.
FT   CONFLICT      1      9       MISSING (IN REF. 2).
FT   CONFLICT    192    192       S -> P (IN REF. 2).
SQ   SEQUENCE   395 AA;  42943 MW;  43B36D0DA22FAFAC CRC64;
```

Query Match 24.8%; Score 535; DB 1; Length 395;
Best Local Similarity 37.2%; Pred. No. 8.1e-23;
Matches 109; Conservative 54; Mismatches 128; Indels 2; Gaps 2;

```
Qy         89 CPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTT 148
Db         65 CANNSDT-LELPASSEALLLGWVPTRLVPAIYGLVVVVGLPANGLALWVLATRVPRLPST 123

Qy        149 VFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISI 208
Db        124 ILLMNLAVADLLLALVLPPRLVYHLRGQRWPFGEAACRVATAALYGHMYGSVLLLAAVSL 183

Qy        209 NRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVH 268
Db        184 DRYLALVHSLRARALRGQRLTTILCLVAWLSAATLVLPLTFHRQTFLLAGSDRMLCHDAL 243

Qy        269 NTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWYVKASLLILVI 328
Db        244 PLAEQTSHWR-PAFICLAVLGCFVPLLAMVLCYGATLRALAANGQRYSHAVRLTALVLFS 302

Qy        329 FTICFAPSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMS 381
Db        303 AVAAFTPSNVLLVLHYSNPSPEAWGNLYGAYVPSLALSTLNSCVDPFIYYYVS 355
```

```
RESULT 9
PAR4_MOUSE
ID   PAR4_MOUSE     STANDARD;      PRT;   396 AA.
AC   O88634;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-
DE   like 3) (Coagulation factor II receptor-like 3).
GN   F2RL3 OR PAR4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98389762; PubMed=9722561;
RA   Kahn M.L., Hammes S.R., Botka C., Coughlin S.R.;
RT   "Gene and locus structure and chromosomal localization of the
RT   protease-activated receptor gene family.";
RL   J. Biol. Chem. 273:23290-23296(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98379991; PubMed=9716134;
RA   Kahn M.L., Zheng Y.-W., Huang W., Bigornia V., Zeng D., Moff S.,
RA   Farese R.V. Jr., Tam C., Coughlin S.R.;
RT   "A dual thrombin receptor system for platelet activation.";
RL   Nature 394:690-694(1998).
CC   -!- FUNCTION: Receptor for activated thrombin or trypsin coupled to G
CC       proteins that stimulate phosphoinositide hydrolysis. May play a
CC       role in platelets activation.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Highly expressed in the spleen. Slight
CC       expression in the heart, lung, skeletal muscle and kidney. No
CC       detectable expression in brain, liver or testis. Also detected in
CC       platelets.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF080215; AAC28861.1; -.
DR   MGD; MGI:1298207; F2rl3.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation.
FT   SIGNAL        1     16       POTENTIAL.
FT   PROPEP       17     59       REMOVED FOR RECEPTOR ACTIVATION
FT                                (BY SIMILARITY).
FT   CHAIN         60    396       PROTEINASE ACTIVATED RECEPTOR 4.
FT   DOMAIN        60     94       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      95    115       1 (POTENTIAL).
FT   DOMAIN       116    120       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     121    141       2 (POTENTIAL).
FT   DOMAIN       142    162       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     163    183       3 (POTENTIAL).
FT   DOMAIN       184    203       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     204    224       4 (POTENTIAL).
FT   DOMAIN       225    255       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     256    276       5 (POTENTIAL).
FT   DOMAIN       277    295       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     296    316       6 (POTENTIAL).
FT   DOMAIN       317    331       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     332    355       7 (POTENTIAL).
FT   DOMAIN       356    396       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       374    377       POLY-SER.
FT   SITE          59     60       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DISULFID     161    240       BY SIMILARITY.
FT   CARBOHYD      68     68       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   396 AA;  42703 MW;  E96636F4804FD3E7 CRC64;

  Query Match             24.6%;  Score 531.5;  DB 1;  Length 396;
  Best Local Similarity   35.7%;  Pred. No. 1.3e-22;
  Matches  112;  Conservative   57;  Mismatches  128;  Indels   17;  Gaps    2;

Qy          84 TVKIKCPEES----------------ASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVG 127
               ||::| |: |                :  | :  ::   |   :||:|:||:| ||  ||
Db          44 TVELKEPKSSDKPNPRGYPGKFCANDSDTLELPASSQALLLGWVSTRLVPALYGLVVAVG 103

Qy         128 VPANAVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRA 187
               :||| : ||:|  |   : :|:  ||||:|| |  :  | ::||||:|  | |||  ||
Db         104 LPANGLALWVLATRVPRLPSTILLTNLAVADSLLALVPPPRLAYHLRGQRWPFGEAACRV 163

Qy         188 TTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPF 247
               |   ||:|| |:|||| :|::||||:||| . | |         | : | :     ||
Db         164 ATAALYGHMYGSVLLLAAVSLDRYLALVHPLRARALRGQRLTTGLCLVAWLSAATLALPL 223

Qy         248 FILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRT 307
               : :| : |   |   |||     | :| ::   || ||  |  :| : :  || | :|
Db         224 TLHRQTFRLAGSDRMLCHDALPLTEQTSHWR-PAFICLAVLGCFVPLLAMGLCYGATLRA 282

Qy         308 LNAYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIALCLGS 367
               | |   |:   :: : |:|      | |||::|::|::|        ||  |: :| | :
Db         283 LAANGQRYSHALRLTALVLFSAVASFTPSNVLLVLHYSNPSPEAWGNLYGAYVPSLALST 342

Qy         368 LNSCLDPFLYFLMS 381
               ||||:|||:|: :|
Db         343 LNSCVDPFIYYYVS 356


RESULT 10
PAR1_MOUSE
ID   PAR1_MOUSE     STANDARD;      PRT;   430 AA.
AC   P30558; P97507;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
GN   F2R OR PAR1 OR CF2R.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Coughlin S.R.;
RL   Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129;
RX   MEDLINE=96379236; PubMed=8784787;
RA   Kahn M.L., Ishii K., Kuo W.L., Piper M., Connolly A., Shi Y.P.,
RA   Wu R., Lin C.C., Coughlin S.R.;
RT   "Conserved structure and adjacent location of the thrombin receptor
RT   and protease-activated receptor 2 genes define a protease-activated
RT   receptor gene cluster.";
RL   Mol. Med. 2:349-357(1996).
RN   [3]
RP   SEQUENCE OF 1-74 FROM N.A.
RC   STRAIN=129/Sv;
RX   MEDLINE=96327649; PubMed=8678993;
RA   Xue J., Jenkins N.A., Gilbert D.J., Copeland N.G., Sadler J.E.;
RT   "Structure and localization of the thrombin receptor gene on mouse
RT   chromosome 13.";
RL   Mamm. Genome 7:625-626(1996).
CC   -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; L03529; AAA40438.1; -.
DR   EMBL; U36757; AAB38308.1; -.
DR   EMBL; U36756; AAB38308.1; JOINED.
DR   EMBL; U55076; AAB00198.1; -.
DR   EMBL; U55075; AAB00198.1; JOINED.
DR   HSSP; P34996; 1DDD.
DR   MGD; MGI:101802; F2r.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     41       REMOVED FOR RECEPTOR ACTIVATION (BY
FT                                SIMILARITY).
FT   CHAIN         42    430       PROTEINASE ACTIVATED RECEPTOR 1.
FT   DOMAIN        42    107       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     108    133       1 (POTENTIAL).
FT   DOMAIN       134    142       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     143    162       2 (POTENTIAL).
FT   DOMAIN       163    181       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     182    203       3 (POTENTIAL).
FT   DOMAIN       204    223       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     224    244       4 (POTENTIAL).
FT   DOMAIN       245    273       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     274    293       5 (POTENTIAL).
FT   DOMAIN       294    316       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     317    339       6 (POTENTIAL).
FT   DOMAIN       340    354       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     355    379       7 (POTENTIAL).
FT   DOMAIN       380    430       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        60     65       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN        88     91       POLY-PRO.
FT   SITE          41     42       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DISULFID     180    259       BY SIMILARITY.
FT   CARBOHYD      67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      80     80       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     255    255       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     162    162       F -> S (IN REF. 1).
FT   CONFLICT     189    189       G -> Y (IN REF. 1).
FT   CONFLICT     223    223       R -> G (IN REF. 1).
FT   CONFLICT     262    262       V -> L (IN REF. 1).
FT   CONFLICT     365    365       S -> T (IN REF. 1).
SQ   SEQUENCE   430 AA;  47790 MW;  395FD64FAE52C9BF CRC64;

  Query Match             24.2%;  Score 522.5;  DB 1;  Length 430;
  Best Local Similarity   30.8%;  Pred. No. 4.1e-22;
  Matches  130;  Conservative   82;  Mismatches  161;  Indels   49;  Gaps   16;

Qy         21 KALIFAAAGLLLLLPTF-CQSGMENDTNNLAKPTLPIKTF--RGAPPNSFEEFPFSALEG 77
              : |:  | || |  |     :  |     :    |:  ::|  |    |:||  |    |
Db          4 RRLLIVALGLSLCGPLLSSRVPMSQPESERTDATVNPRSFFLRNPSENTFELVPLGDEEE 63

Qy         78 --------WTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVP 129
                      |  : : |  |  :     :      |||||   |  :|::| :||:| :|
Db         64 EEKNESVLLEGRAVYLNISLPPHTPPPPFISEDASGYLTSPWLTLFMPSVYTIVFIVSLP 123

Qy        130 AN--AVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRA 187
              |  |: :::|  : :     |:  :||:|| ||   |||||:|: :| :| ||  :||
Db        124 LNVLAIAVFVLRMKVKK-PAVVYMLHLAMADVLFVSVLPFKISYYFSGTDWQFGSGMCRF 182

Qy        188 TTVIFYGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLYM 244
              |  |||||| ||:|:  |||:|:||:|:|    ::| | :  :   || ::|    : :
Db        183 ATAAFYGNMYASIMLMTVISIDRFLAVVYPIQSLSWRTLGRANF---TCVVIWVMAIMGV 239

Qy        245 LPFFILKQEYYLVQPDITTCHDV--HNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYA 302
              :|  : :|   :   :|||||||   | :    |  ||| : :   ||:| ::   ||
Db        240 VPLLLKEQTTRVPGLNITTCHDVLSENLMQG---FYSYYFSAFSAIFFLVPLIVSTVCYT 296

Qy        303 AIIRTLN-------AYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYY----NN 351
              :||| |:       :   | |:    |  :  || :|| |:|::||:|:   | :
Db        297 SIIRCLSSSAVANRSKKSRALF---LSAAVFCIFIVCFGPTNVLLIVH---YLFLSDSPG 350

Qy        352 TDGLYFIYLIALCLGSLNSCLDPFL-YFLMSKTRNHSTAYL----TKXXNDLREQGQ--P 404
              |:  || ||: :|: |:: |:|| : |:  |: : |  : |    :   |     ||  |
Db        351 TEAAYFAYLLCVCVSSVSCCIDPLIYYYASSECQRHLYSILCCKESSDPNSCNSTGQLMP 410

Qy        405 SQ 406
              |:
Db        411 SK 412


RESULT 11
PAR4_HUMAN
ID   PAR4_HUMAN     STANDARD;      PRT;   385 AA.
AC   Q96RI0; O76067;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-
DE   like 3) (Coagulation factor II receptor-like 3).
GN   F2RL3 OR PAR4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98389762; PubMed=9722561;
RA   Kahn M.L., Hammes S.R., Botka C., Coughlin S.R.;
RT   "Gene and locus structure and chromosomal localization of the
RT   protease-activated receptor gene family.";
RL   J. Biol. Chem. 273:23290-23296(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98379991; PubMed=9716134;
RA   Kahn M.L., Zheng Y.-W., Huang W., Bigornia V., Zeng D., Moff S.,
RA   Farese R.V. Jr., Tam C., Coughlin S.R.;
RT   "A dual thrombin receptor system for platelet activation.";
RL   Nature 394:690-694(1998).
RN   [3]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-47 AND ARG-68.
RC   TISSUE=Lymphoma;
RX   MEDLINE=98283984; PubMed=9618465;
RA   Xu W.-F., Andersen H., Whitmore T.E., Presnell S.R., Yee D.P.,
RA   Ching A., Gilbert T., Davie E.W., Foster D.C.;
RT   "Cloning and characterization of human protease-activated receptor
RT   4.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:6642-6646(1998).
RN   [4]
RP   SEQUENCE FROM N.A., AND VARIANTS VAL-296 AND LEU-310.
RA   Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   FUNCTION.
RX   MEDLINE=99178892; PubMed=10079109;
RA   Kahn M.L., Nakanishi-Matsui M., Shapiro M.J., Ishihara H.,
RA   Coughlin S.R.;
RT   "Protease-activated receptors 1 and 4 mediate activation of human
RT   platelets by thrombin.";
RL   J. Clin. Invest. 103:879-887(1999).
CC   -!- FUNCTION: Receptor for activated thrombin or trypsin coupled to G
CC       proteins that stimulate phosphoinositide hydrolysis. May play a
CC       role in platelets activation.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in lung,
CC       pancreas, thyroid, testis and small intestine. Not expressed in
CC       brain, kidney, spinal cord and peripheral blood leukocytes. Also
CC       detected in platelets.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF080214; AAC28860.1; -.
DR   EMBL; AF055917; AAC25699.1; -.
DR   EMBL; AF384819; AAK61908.1; -.
DR   Genew; HGNC:3540; F2RL3.
DR   MIM; 602779; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation; Polymorphism.
FT   SIGNAL        1     17       POTENTIAL.
FT   PROPEP       18     47       REMOVED FOR RECEPTOR ACTIVATION
FT                                (BY SIMILARITY).
FT   CHAIN        48    385       PROTEINASE ACTIVATED RECEPTOR 4.
FT   DOMAIN       48     82       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     83    103       1 (POTENTIAL).
FT   DOMAIN      104    108       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    109    129       2 (POTENTIAL).
FT   DOMAIN      130    151       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    152    172       3 (POTENTIAL).
FT   DOMAIN      173    192       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    193    213       4 (POTENTIAL).
FT   DOMAIN      214    247       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    248    268       5 (POTENTIAL).
FT   DOMAIN      269    283       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    284    304       6 (POTENTIAL).
FT   DOMAIN      305    319       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    320    343       7 (POTENTIAL).
FT   DOMAIN      344    385       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      301    305       POLY-LEU.
FT   SITE         47     48       CLEAVAGE (BY THROMBIN OR TRYPSIN) (BY
FT                                SIMILARITY).
FT   DISULFID    149    228       BY SIMILARITY.
FT   CARBOHYD     56     56       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     296    296       F -> V.
FT                                /FTId=VAR_012852.
FT   VARIANT     310    310       P -> L.
FT                                /FTId=VAR_012853.
FT   MUTAGEN      47     47       R->A: NO PROTEOLYTIC CLEAVAGE (BY
FT                                THROMBIN OR TRYPSIN).
FT   MUTAGEN      68     68       R->A: NO EFFECT ON RECEPTOR ACTIVATION.
FT   CONFLICT    120    120       T -> A (IN REF. 4).
SQ   SEQUENCE   385 AA;  41162 MW;  FACA62B40D5756AE CRC64;

  Query Match           24.2%;  Score 521.5;  DB 1;  Length 385;
  Best Local Similarity  33.8%;  Pred. No. 4.3e-22;
  Matches  120;  Conservative   62;  Mismatches  158;  Indels   15;  Gaps    5;

Qy         29 GLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPNSFEEFP--FSALEGWTGATITVK 86
              | ||| |      :   |      |  :    |:     :  |    |  |: |
Db          3 GRLLLWPLVLGFSLSGGTQ-----TPSVYDESGSTGGGDDSTPSILPAPRGYPGQV---- 53

Qy         87 IKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC 146
              |  :| : | : :::   |   : |:|:||:| || |||:||| : ||:|  :   :
Db         54 --CANDSDT-LELPDSSRALLLGWVPTRLVPALYGLVLVVGLPANGLALWVLATQAPRLP 110

Qy        147 TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACI 206
             :|:   ||| || |  : || :||||| |  | |||  ||  |   ||:|| |:|||| :
Db        111 STMLLMNLATADLLLALALPPRIAYHLRGQRWPFGEAACRLATAALYGHMYGSVLLLAAV 170

Qy        207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHD 266
             |::||||:|||   | |     ||  |   |       ||  : :| : | : |   |||
Db        171 SLDRYLALVHPLRARALRGRRLALGLCMAAWLMAAALALPLTLQRQTFRLARSDRVLCHD 230

Qy        267 VHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWYVKASLLIL 326
                  :| :|   |  ||  | :| : :: || | : || |   |:   :: : ::|
Db        231 ALPLDAQASHWQ-PAFTCLALLGCFLPLLAMLLCYGATLHTLAASGRRYGHALRLTAVVL 289

Qy        327 VIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMS 381
                 | |||::|::|:::   :    ||  |: :| | :||||:|||:|: :|
Db        290 ASAVAFFVPSNLLLLLHYSDPSPSAWGNLYGAYVPSLALSTLNSCVDPFIYYYVS 344


RESULT 12
PAR1_PAPHA
ID   PAR1_PAPHA     STANDARD;      PRT;   425 AA.
AC   P56488;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
GN   F2R OR PAR1 OR BTHR12.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shoji M., Hayzer D.J., Hanson S.R.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF028727; AAB84191.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation.
FT   SIGNAL        1     26       POTENTIAL.
FT   PROPEP       27     41       REMOVED FOR RECEPTOR ACTIVATION.
FT   CHAIN        42    425       PROTEINASE ACTIVATED RECEPTOR 1.
FT   DOMAIN       42    102       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    103    128       1 (POTENTIAL).
FT   DOMAIN      129    137       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    138    157       2 (POTENTIAL).
FT   DOMAIN      158    176       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    177    198       3 (POTENTIAL).
FT   DOMAIN      199    218       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    219    239       4 (POTENTIAL).
FT   DOMAIN      240    268       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    269    288       5 (POTENTIAL).
FT   DOMAIN      289    311       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    312    334       6 (POTENTIAL).
FT   DOMAIN      335    350       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    351    374       7 (POTENTIAL).
FT   DOMAIN      375    425       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       57     60       ASP/GLU-RICH (ACIDIC).
FT   SITE         41     42       CLEAVAGE (BY THROMBIN).
FT   DISULFID    175    254       BY SIMILARITY.
FT   CARBOHYD     35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     75     75       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    250    250       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    259    259       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   425 AA;  47253 MW;  118FC5FB39D4DE2C CRC64;

  Query Match             24.1%;  Score 520.5;  DB 1;  Length 425;
  Best Local Similarity   32.9%;  Pred. No. 5.2e-22;
  Matches  125;  Conservative   75;  Mismatches  153;  Indels   27;  Gaps   11;

Qy         21 KALIFAAAGLLLLLPTF-CQSGMENDTNNLAKPTLPIKTFRGAPPN-SFEEFPFSALEGW 78
              : |:  || | | |    ::      :     ||  ::|    ||  :| |     :
Db          4 RRLLLVAACLCLCGPLLSARTRARRPASKATNATLDPRSFLLRNPNDKYEPFWEDEEKNE 63

Qy         79 TGAT----ITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPAN--A 132
              :| |    :::     | :      :     ||||||  |  :|::|  |||| :| |  |
Db         64 SGLTEYRLVSINKSSPLQKPLPAFISEDASGYLTSSWLTLFVPSVYTGVFVVSLPVNIMA 123

Qy        133 VTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIF 192
              : :::|  : :     |:  :|| || ||   |||||:|:|:|::| ||  |||  |  |
Db        124 IVVFILKMKVKK-PAVVYMLHLATADVLFVSVLPFKISYYLSGSDWQFGSELCRFVTAAF 182

Qy        193 YGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLYMLPFFI 249
              | ||| ||||:  |||:|:||:|:|    ::| | : ::   ||  :||     ::|  :
Db        183 YCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASF---TCLAIWALAIAGVVPLLL 239

Qy        250 LKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLN 309
              :|   :   :|||||| |       :  ||| : :   | :| ::   || :||| |:
Db        240 KEQTIQVPGLNITTCHDVLNETLLEG-YYAYYFSAFSAVFFFVPLIISTVCYVSIIRCLS 298

Qy        310 -------AYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANY-YYNNTDGLYFIYLI 361
                    :   | |:    |  :  || ||| |:||:|| |::   : : |:  || ||:
Db        299 SSTVANRSKKSRALF---LSAAVFCIFIICFGPTNILLIAHYSFLSHTSTTEAAYFAYLL 355

Qy        362 ALCLGSLNSCLDPFLYFLMS 381
              :|: |::|:|| :|:  |
Db        356 CVCVSSISCCIDPLIYYYAS 375


RESULT 13
PAR1_HUMAN
ID   PAR1_HUMAN     STANDARD;      PRT;   425 AA.
AC   P25116; Q96RF7; Q9BUN4;
DT   01-MAY-1992 (Rel. 22, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor)
DE   (Coagulation factor II receptor).
GN   F2R OR PAR1 OR TR OR CF2R.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91168254; PubMed=1672265;
RA   Vu T.-K.H., Hung D.T., Wheaton V.I., Coughlin S.R.;
RT   "Molecular cloning of a functional thrombin receptor reveals a novel
RT   proteolytic mechanism of receptor activation.";
RL   Cell 64:1057-1068(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   FUNCTION.
RX   MEDLINE=99178892; PubMed=10079109;
RA   Kahn M.L., Nakanishi-Matsui M., Shapiro M.J., Ishihara H.,
RA   Coughlin S.R.;
RT   "Protease-activated receptors 1 and 4 mediate activation of human
RT   platelets by thrombin.";
RL   J. Clin. Invest. 103:879-887(1999).
CC   -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis. May play a
CC       role in platelets activation and in vascular development.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: PLATELETS AND VASCULAR ENDOTHELIAL CELLS.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- PTM: Phosphorylated; probably mediating desenzitisation prior to
CC       the uncoupling and internalization of the receptor.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M62424; AAA36743.1; -.
DR   EMBL; AF391809; AAK69768.1; -.
DR   EMBL; BC002464; AAH02464.1; -.
DR   PIR; A37912; A37912.
DR   HSSP; P34996; 1DDD.
DR   Genew; HGNC:3537; F2R.
DR   MIM; 187930; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation; Phosphorylation.
FT   SIGNAL       1     26       POTENTIAL.
FT   PROPEP      27     41       REMOVED FOR RECEPTOR ACTIVATION.
FT   CHAIN       42    425       PROTEINASE ACTIVATED RECEPTOR 1.
FT   DOMAIN      42    102       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   103    128       1 (POTENTIAL).
FT   DOMAIN     129    137       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   138    157       2 (POTENTIAL).
FT   DOMAIN     158    176       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   177    198       3 (POTENTIAL).
FT   DOMAIN     199    218       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   219    239       4 (POTENTIAL).
FT   DOMAIN     240    268       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   269    288       5 (POTENTIAL).
FT   DOMAIN     289    311       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   312    334       6 (POTENTIAL).
FT   DOMAIN     335    350       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   351    374       7 (POTENTIAL).
FT   DOMAIN     375    425       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      57     60       ASP/GLU-RICH (ACIDIC).
FT   SITE        41     42       CLEAVAGE (BY THROMBIN).
FT   DISULFID   175    254       BY SIMILARITY.
FT   CARBOHYD    35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    75     75       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   250    250       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   259    259       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   238    238       L -> V (IN REF. 1).
FT   CONFLICT   335    335       A -> V (IN REF. 3).
FT   CONFLICT   364    364       C -> S (IN REF. 1).
SQ   SEQUENCE   425 AA;  47440 MW;  41B742A99EEC96AB CRC64;

  Query Match             24.1%;  Score 519;  DB 1;  Length 425;
  Best Local Similarity   32.6%;  Pred. No. 6.3e-22;
  Matches  126;  Conservative   76;  Mismatches  154;  Indels   30;  Gaps   11;

Qy         14 GTQVIKMKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPN-SFEEFPF 72
              | : : : |  |:  | ||   |  :      ||     ||  ::|    ||  :| |
```

```
Db          2 GPRRLLLVAACFSLCGPLLSARTRARRPESKATN----ATLDPRSFLLRNPNDKYEPFWE 57

Qy         73 SALEGWTGAT----ITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGV 128
                 :  :| |    :::    | :      :    ||||||  |  :|::|  |||| :
Db         58 DEEKNESGLTEYRLVSINKSSPLQKQLPAFISEDASGYLTSSWLTLFVPSVYTGVFVVSL 117

Qy        129 PAN--AVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCR 186
              | |  |: :::|  : :     |:  :|| || ||   |||||:|: :|::| ||  |||
Db        118 PLNIMAIVVFILKMKVKK-PAVVYMLHLATADVLFVSVLPFKISYYFSGSDWQFGSELCR 176

Qy        187 ATTVIFYGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
                | || ||| ||||:  |||:|:||:|:|    ::| |.: ::   ||  :||
Db        177 FVTAAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASF---TCLAIWALAIAG 233

Qy        244 MLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAA 303
              ::|  : :|   :   :|||||| |       :  ||| : :  | :| ::   || :
Db        234 VVPLLLKEQTIQVPGLNITTCHDVLNETLLEG-YYAYYFSAFSAVFFFVPLIISTVCYVS 292

Qy        304 IIRTLN-------AYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANY-YYNNTDGL 355
              ||| |:       :   | |:    |  :  || ||| |:|::|| |::   : : |:
Db        293 IIRCLSSSAVANRSKKSRALF---LSAAVFCIFIICFGPTNVLLIAHYSFLSHTSTTEAA 349

Qy        356 YFIYLIALCLGSLNSCLDPFLYFLMS 381
              || ||: :|: |:: |:|| :|:  |
Db        350 YFAYLLCVCVSSISCCIDPLIYYYAS 375


RESULT 14
PAR1_RAT
ID   PAR1_RAT       STANDARD;      PRT;   432 AA.
AC   P26824;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
GN   F2R OR PAR1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aortic smooth muscle;
RX   MEDLINE=92381002; PubMed=1324917;
RA   Zhong C., Hayzer D.J., Corsen M.A., Wick K., Runge M.S.;
RT   "Molecular cloning of the rat vascular smooth muscle thrombin
RT   receptor. Evidence for in vitro regulation by basic fibroblast growth
RT   factor.";
RL   J. Biol. Chem. 267:16975-16979(1992).
CC   -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M81642; AAA42274.1; -.
DR   PIR; A43448; A43448.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     45       REMOVED FOR RECEPTOR ACTIVATION (BY
FT                                SIMILARITY).
FT   CHAIN        46    432       PROTEINASE ACTIVATED RECEPTOR 1.
FT   DOMAIN       46    109       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    110    135       1 (POTENTIAL).
FT   DOMAIN      136    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    145    164       2 (POTENTIAL).
FT   DOMAIN      165    183       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    184    205       3 (POTENTIAL).
FT   DOMAIN      206    225       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    226    246       4 (POTENTIAL).
FT   DOMAIN      247    275       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    276    295       5 (POTENTIAL).
FT   DOMAIN      296    318       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    319    341       6 (POTENTIAL).
FT   DOMAIN      342    357       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    358    381       7 (POTENTIAL).
FT   DOMAIN      382    432       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       64     68       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN       87     93       POLY-PRO.
FT   SITE         45     46       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DISULFID    182    261       BY SIMILARITY.
FT   CARBOHYD     69     69       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    257    257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    266    266       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   432 AA;  48280 MW;  DD032B97ABA4A605 CRC64;

  Query Match          23.4%;  Score 505.5;  DB 1;  Length 432;
  Best Local Similarity  31.2%;  Pred. No. 3.4e-21;
  Matches  134;  Conservative   71;  Mismatches  163;  Indels   61;  Gaps   15;

Qy         21 KALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKT-------FRGAPPNSFEEFPFS 73
                : |:  | || |  |          :      |  |       |    ::||:||
Db          4 RRLLLVAVGLSLCGPLLSSRVPMRQPESERMYATPYATPNPRSFFLRNPSEDTFEQFPLG 63

Qy         74 ---------ALEGWTGATITVKIKCP--------EESASHLHVKNATMGYLTSSLSTKLI 116
                       |||   |   | : |         | ||          |||||   |  |
Db         64 DEEEKNESIPLEG--RAVYLNKSRFPPMPPPPFISEDAS---------GYLTSPWLTLFI 112

Qy        117 PAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFY-TNLAIADFLFCVTLPFKIAYHLNG 175
              |::|  ||:| :| | : : :  || :     | |  :||:|| ||   |||||:|: :|
Db        113 PSVYTFVFIVSLPLNILAIAVFVFRMKVKKPAVVYMLHLAMADVLFVSVLPFKISYYFSG 172

Qy        176 NNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVT 232
               :| ||  :||  |   | ||| ||:|:  |||:|:||:|:|    ::| | :  :   |
Db        173 TDWQFGSGMCRFATAACYCNMYASIMLMTVISIDRFLAVVYPIQSLSWRTLGRANF---T 229

Qy        233 CGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLI 292
              . | ::|    : ::|  : :|   :   :|||||| |      .  | ||| : :   ||:
Db        230 CVVIWVMAIMGVVPLLLKEQTTQVPGLNITTCHDVLNETLLHG-FYSYYFSAFSAIFFLV 288

Qy        293 PFVLIIYCYAAIIRTLN-------AYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHA 345
              | ::   || :||| |:       :   | |:    |  :  || :|| |:|::||:|:
Db        289 PLIISTVCYTSIIRCLSSSAVANRSKKSRALF---LSAAVFCIFIVCFGPTNVLLIVHYL 345

Qy        346 NYYYN-NTDGLYFIYLIALCLGSLNSCLDPFL-YFLMSKTRNHSTAYL----TKXXNDLR 399
                  :  |:  || ||: :|: |: ||:|| : |:  |: : |  : |    |
Db        346 LLSDSPGTETAYFAYLLCVCVTSVASCIDPLIYYYASSECQKHLYSILCCRESSDSNSCN 405

Qy        400 EQGQ--PSQ 406
                ||  ||:
Db        406 STGQLMPSK 414


RESULT 15
PAR1_CRILO
ID   PAR1_CRILO     STANDARD;      PRT;   428 AA.
AC   Q00991; Q60461;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
```

```
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
GN   F2R OR PAR1.
OS   Cricetulus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10030;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=91348247; PubMed=1652467;
RA   Rasmussen U.B., Vouret-Craviari V., Jallat S., Schlesinger Y.,
RA   Pages G., Pavirani A., Lecocq J.P., Pouyssegur J.,
RA   Obberghen-Schilling E.;
RT   "cDNA cloning and expression of a hamster alpha-thrombin receptor
RT   coupled to Ca2+ mobilization.";
RL   FEBS Lett. 288:123-128(1991).
RN   [2]
RP   SEQUENCE OF 42-428 FROM N.A.
RX   MEDLINE=96028007; PubMed=7488069;
RA   Hartmann T., Grace M.B., Buzard G.S., Ruoss S.J.;
RT   "Thrombin receptor polymorphism in Chinese hamster.";
RL   Biochem. Biophys. Res. Commun. 215:974-980(1995).
CC   -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC       G proteins that stimulate phosphoinositide hydrolysis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC       functions as a tethered ligand.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61958; CAA43957.1; -.
DR   EMBL; U34047; AAA86747.1; -.
DR   PIR; S17148; S17148.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Blood coagulation; Polymorphism.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     41       REMOVED FOR RECEPTOR ACTIVATION (BY
FT                                SIMILARITY).
FT   CHAIN        42    428       PROTEINASE ACTIVATED RECEPTOR 1.
FT   DOMAIN       42    105       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    106    131       1 (POTENTIAL).
FT   DOMAIN      132    140       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    141    160       2 (POTENTIAL).
FT   DOMAIN      161    179       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    180    201       3 (POTENTIAL).
FT   DOMAIN      202    221       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    222    242       4 (POTENTIAL).
FT   DOMAIN      243    271       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    272    291       5 (POTENTIAL).
FT   DOMAIN      292    314       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    315    337       6 (POTENTIAL).
FT   DOMAIN      338    352       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    353    377       7 (POTENTIAL).
FT   DOMAIN      378    428       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       60     63       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      301    304       POLY-SER.
FT   SITE         41     42       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DISULFID    178    257       BY SIMILARITY.
FT   CARBOHYD     65     65       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     78     78       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    253    253       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    262    262       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      84     84       MISSING (IN AN ALLELE).
FT   CONFLICT    384    384       H -> T (IN REF. 1).
SQ   SEQUENCE   428 AA;  47602 MW;  E5492AB2CD877E2F CRC64;

  Query Match             23.2%;  Score 501;  DB 1;  Length 428;
  Best Local Similarity   32.1%;  Pred. No. 6e-21;
  Matches  126;  Conservative   73;  Mismatches  140;  Indels   54;  Gaps   13;

Qy         23 LIFAAAGLLLLLPTFC--------QSGMENDTNNLAKPTLPIKTF-RGAPPNSFEEFPFS 73
              |:  |||| |  |            :| | : | |       |   | |    |:||  |
Db          6 LLLVAAGLSLCGPLLSSRVPVRQPESEMTDATVN------PRSFFLRNPGENTFELIPLG 59

Qy         74 ALEGWTGATITVKIKCPEESASHLH------------VKNATMGYLTSSLSTKLIPAIYL 121
                |    :|:      ||  | :|:           :     |||||      ||::|
Db         60 DEEEKNESTL------PEGRAIYLNKSHSPPAPLAPFISEDASGYLTSPWLRLFIPSVYT 113

Qy        122 LVFVVGVPAN--AVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV 179
               |||| :| |  |: :::|  : :     |:  :||:|| ||   || ||:|: :|::|
Db        114 FVFVVSLPLNILAIAVFVLKMKVKK-PAVVYMLHLAMADVLFVSVLPLKISYYFSGSDWQ 172

Qy        180 FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLV 236
               ||  :||  |  || ||| ||:|:  |||:|:||:|:|    ::| | :  :   || ::
Db        173 FGSGMCRFATAAFYCNMYASIMLMTVISIDRFLAVVYPIQSLSWRTLGRANF---TCLVI 229

Qy        237 WATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVL 296
               |    : ::|  : :|   :   :|||||| |       | ||| : :   ||:| ::
Db        230 WVMAIMGVVPLLLKEQTTRVPGLNITTCHDVLNETLLQG-FYSYYFSAFSAVFFLVPLII 288

Qy        297 IIYCYAAIIRTLNAYD-------HRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYY 349
                 || :||| |::          | |:    |  :  :| :|| |:|::||:|:
Db        289 STICYMSIIRCLSSSSVANRSKKSRALF---LSAAVFCVFIVCFGPTNVLLIMHYLLLSD 345

Qy        350 N-NTDGLYFIYLIALCLGSLNSCLDPFLYFLMS 381
              :  |:  || ||: :|: |:: |:|| :|:  |
Db        346 SPATEKAYFAYLLCVCVSSVSCCIDPLIYYYAS 378


RESULT 16
P2Y8_XENLA
ID   P2Y8_XENLA     STANDARD;      PRT;   537 AA.
AC   P79928;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   P2Y purinoceptor 8 (P2Y8).
GN   P2RY8.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Neural plate;
RX   MEDLINE=97284734; PubMed=9139711;
RA   Bogdanov Y.D., Dale L., King B.F., Whittock N., Burnstock G.;
RT   "Early expression of a novel nucleotide receptor in the neural plate
RT   of Xenopus embryos.";
RL   J. Biol. Chem. 272:12583-12590(1997).
CC   -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ATP, UTP, CTP, GTP AND ITP.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC       MAY PLAY A KEY ROLE IN THE EARLY DEVELOPMENT OF NEURAL TISSUE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X99953; CAA68213.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     49       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     50     70       1 (POTENTIAL).
FT   DOMAIN       71     79       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     80    100       2 (POTENTIAL).
FT   DOMAIN      101    118       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    119    139       3 (POTENTIAL).
FT   DOMAIN      140    161       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    162    182       4 (POTENTIAL).
FT   DOMAIN      183    210       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    211    231       5 (POTENTIAL).
FT   DOMAIN      232    254       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    255    275       6 (POTENTIAL).
FT   DOMAIN      276    292       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    293    316       7 (POTENTIAL).
FT   DOMAIN      317    537       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    116    193       BY SIMILARITY.
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     29     29       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   537 AA;  62024 MW;  B2CF24812F3C19F2 CRC64;

  Query Match             19.8%;  Score 428;  DB 1;  Length 537;
  Best Local Similarity   36.0%;  Pred. No. 6.4e-17;
  Matches  105;  Conservative   41;  Mismatches  102;  Indels   44;  Gaps   12;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRS-ICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              |:|  |  ||:||:| |   :|:   : |    |||:  |||::| |: ::||  : |:
Db         45 LLPVSYSAVFMVGLPLNIAAMWIFIAKMRPWNPTTVYMFNLALSDTLYVLSLPTLVYYYA 104

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFT--YRGLPKHTYALV 231
              : ||| ||||||:    :|| |:| ||| | |||::||  : || |   |   || | |
Db        105 DKNNWPFGEVLCKLVRFLFYANLYSSILFLTCISVHRYRGVCHPITSLRRMNAKHAY--V 162

Qy        232 TCGLVWATVFLYMLPFFILKQEYYLVQPDI--TTCHDVHNTCESSSPFQLYYFISLAFFG 289
               | ||| :| | ::|  |    :  | | :  | |||          |  |   | |
Db        163 ICALVWLSVTLCLVPNLI----FVTVSPKVKNTICHDT----TRPEDFARYVEYSTAIMC 214

Qy        290 FL--IPFVLIIYCYAAIIR------------TLNAYDHRWLWYVKASLLILVIFTICFAP 335
               |  || ::|  ||  : |             || :|  |    :|  : ::: | ||| |
Db        215 LLFGIPCLIIAGCYGLMTRELMKPIVSGNQQTLPSYKKR---SIKTIIFVMIAFAICFMP 271

Qy        336 SNIILIIHHANYYYNNTDGL--YFI------YLIALCLGSLNSCLDPFLYFL 379
              :|   :    |||    |:  | :       | :   | | |||:|| ||||
Db        272 FHITRTL----YYYARLLGIKCYALNVINVTYKVTRPLASANSCIDPILYFL 319


RESULT 17
P2YR_MELGA
ID   P2YR_MELGA     STANDARD;      PRT;   362 AA.
AC   P49652;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor) (6H1
DE   orphan receptor).
GN   P2RY1.
OS   Meleagris gallopavo (Common turkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX   NCBI_TaxID=9103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=94335907; PubMed=8058061;
RA   Filtz T.M., Li Q., Boyer J.L., Nicholas R.A., Harden T.K.;
RT   "Expression of a cloned P2Y purinergic receptor that couples to
RT   phospholipase C.";
RL   Mol. Pharmacol. 46:8-14(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97382456; PubMed=9240460;
RA   Li Q., Schachter J.B., Harden T.K., Nicholas R.A.;
RT   "The 6H1 orphan receptor, claimed to be the p2y5 receptor, does not
RT   mediate nucleotide-promoted second messenger responses.";
RL   Biochem. Biophys. Res. Commun. 236:455-460(1997).
CC   -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC       ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN
CC       INSENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT
CC       ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: MAINLY FOUND IN BLOOD, BRAIN, AND LUNG. TO A
CC       LESSER EXTENT IN STOMACH, GUT, AND SKELETAL MUSCLE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U09842; AAA18784.1; -.
DR   EMBL; AF012103; AAB65428.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     41       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     42     63       1 (POTENTIAL).
FT   DOMAIN       64     76       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     77     98       2 (POTENTIAL).
FT   DOMAIN       99    115       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    116    136       3 (POTENTIAL).
FT   DOMAIN      137    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    177       4 (POTENTIAL).
FT   DOMAIN      178    207       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    208    227       5 (POTENTIAL).
FT   DOMAIN      228    254       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    255    274       6 (POTENTIAL).
FT   DOMAIN      275    292       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    293    317       7 (POTENTIAL).
FT   DOMAIN      318    362       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    113    191       BY SIMILARITY.
FT   CARBOHYD     11     11       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    102    102       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    186    186       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   362 AA;  41180 MW;  3E128AB9EB64349C CRC64;

  Query Match             19.7%;  Score 426;  DB 1;  Length 362;
  Best Local Similarity   30.4%;  Pred. No. 6.1e-17;
  Matches  101;  Conservative   59;  Mismatches  132;  Indels   40;  Gaps    9;

Qy        105 GYLTSSLSTK-----------LIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYT 152
              |:   : |||            :| :|:|||: |·  |:| :||  |  |     :|:
Db         21 GWAAGNASTKCSLTKTGFQFYYLPTVYILVFITGFLGNSVAIWMFVFHMRPWSGISVYMF 80

Qy        153 NLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYL 212
              |||:|||: :|||  | |: |  :|:||:|:|:    ||: |:| ||| | |||::||
Db         81 NLALADFLYVLTLPALIFYYFNKTDWIFGDVMCKLQRFIFHVNLYGSILFLTCISVHRYT 140
```

```
Qy        213 AIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY----LVQPDITTCHDVH 268
              :|||    |  |   |:    |||| |   : |       :|    : :    ||:|
Db        141 GVVHPLKSLGRLKKKNAVYVSSLVWALVVAVIAPIL-----FYSGTGVRRNKTITCYD-- 193

Qy        269 NTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKA 321
              | :        | :    | | |||::|: ||  |:: |   |        : :: |
Db        194 TTADEYLRSYFVYSMCTTVFMFCIPFIVILGCYGLIVKALIYKDLDNSPLRRKSIYLV-- 251

Qy        322 SLLILVIFTICFAPSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDP 374
              :::| :| : : | :::  ::  |   :         | :|  | :   | |||||:||
Db        252 -IIVLTVFAVSYLPFHVMKTLNLRARLDFQTPQMCAFNDKVYATYQVTRGLASLNSCVDP 310

Qy        375 FLYFLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
              ||||   |     :  |:  :   |    |:
Db        311 ILYFLAGDTFRRRLSRATRKSSRRSEPNVQSK 342


RESULT 18
P2YR_CHICK
ID   P2YR_CHICK     STANDARD;      PRT;   362 AA.
AC   P34996;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=93285340; PubMed=8508924;
RA   Webb T.E., Simon J., Krishek B.J., Bateson A.N., Smart T.G.,
RA   King B.F., Burnstock G., Barnard E.A.;
RT   "Cloning and functional expression of a brain G-protein-coupled ATP
RT   receptor.";
RL   FEBS Lett. 324:219-225(1993).
RN   [2]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE=97026278; PubMed=8872457;
RA   van Rhee A.M., Fischer B., van Galen P.J.M., Jacobson K.A.;
RT   "Modelling the P2Y purinoceptor using rhodopsin as template.";
RL   Drug Des. Discov. 13:133-140(1995).
CC   -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC       ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN
CC       INSENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT
CC       ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, GASTROINTESTINAL TRACT,
CC       SPLEEN AND LEG MUSCLE. IS NOT DETECTED IN THE HEART, LIVER,
CC       STOMACH, LUNG AND KIDNEY.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X73268; CAA51716.1; -.
DR   PIR; S33733; S33733.
DR   PDB; 1DDD; 11-JUL-96.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; 3D-structure.
FT   DOMAIN       1     41       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    42     63       1 (POTENTIAL).
FT   DOMAIN      64     76       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    77     98       2 (POTENTIAL).
FT   DOMAIN      99    115       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   116    136       3 (POTENTIAL).
FT   DOMAIN     137    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   156    177       4 (POTENTIAL).
FT   DOMAIN     178    207       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   208    227       5 (POTENTIAL).
FT   DOMAIN     228    254       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   255    274       6 (POTENTIAL).
FT   DOMAIN     275    292       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   293    317       7 (POTENTIAL).
FT   DOMAIN     318    362       CYTOPLASMIC (POTENTIAL).
FT   DISULFID   113    191       BY SIMILARITY.
FT   CARBOHYD    11     11       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   102    102       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   186    186       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   362 AA;  41194 MW;  A806C88FB9514761 CRC64;

  Query Match             19.7%;  Score 424;  DB 1;  Length 362;
  Best Local Similarity   28.9%;  Pred. No. 7.8e-17;
  Matches  101;  Conservative   59;  Mismatches  143;  Indels   46;  Gaps    9;

Qy         77 GWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLW 136
              ||      | |    :      :                  :| :|:|||: |   |:| :|
Db         21 GWAAGNATTKCSLTKTGFQFYY----------------LPTVYILVFITGFLGNSVAIW 63

Qy        137 MLFFRTRSIC-TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGN 195
              |  |  |     :|:  |||:||||: :|||  | |: |  :|:||:|:|:    ||: |
Db         64 MFVFHMRPWSGISVYMFNLALADFLYVLTLPALIFYYFNKTDWIFGDVMCKLQRFIFHVN 123

Qy        196 MYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY 255
              :| ||| | |||::||  :|||     |  |    |:    |||| |   : |      :|
Db        124 LYGSILFLTCISVHRYTGVVHPLKSLGRLKKKNAVYVSSLVWALVVAVIAPIL-----FY 178

Qy        256 ----LVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAY 311
                  : :    ||:|   | :        | :    | | |||::|: ||  |:: | ·
Db        179 SGTGVRRNKTITCYD--TTADEYLRSYFVYSMCTTVFMFCIPFIVILGCYGLIVKALIYK 236

Qy        312 D-------HRWLWYVKASLLILVIFTICFAPSNIILIIH-HANYYYNN------TDGLYF 357
              |       : :: |   :::| :| : : | :::  ::  |   :          | :|
Db        237 DLDNSPLRRKSIYLV---IIVLTVFAVSYLPFHVMKTLNLRARLDFQTPQMCAFNDKVYA 293

Qy        358 IYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
              | :   | |||||:|| ||||   |      :  |:  :   |    |:
Db        294 TYQVTRGLASLNSCVDPILYFLAGDTFRRRLSRATRKSSRRSEPNVQSK 342


RESULT 19
EBI2_HUMAN
ID   EBI2_HUMAN     STANDARD;      PRT;   361 AA.
AC   P32249;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   EBV-induced G protein-coupled receptor 2 (EBI2).
GN   EBI2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93188173; PubMed=8383238;
RA   Birkenbach M.P., Josefsen K., Yalamanchili R.R., Lenoir G.M.,
RA   Kieff E.;
RT   "Epstein-Barr virus-induced genes: first lymphocyte-specific G
RT   protein-coupled peptide receptors.";
RL   J. Virol. 67:2209-2220(1993).
CC   -!- FUNCTION: ORPHAN RECEPTOR. PROBABLE MEDIATOR OF EBV EFFECTS ON B
```

```
CC       LYMPHOCYTES OR OF NORMAL LYMPHOCYTE FUNCTIONS.
CC   -!- TISSUE SPECIFICITY: B-LYMPHOCYTE CELL LINES.
CC   -!- INDUCTION: BY EBV.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-5 IS THE INITIATOR.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L08177; AAA35924.1; -.
DR   PIR; B45680; B45680.
DR   Genew; HGNC:3128; EBI2.
DR   MIM; 605741; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     31       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     32     57       1 (POTENTIAL).
FT   DOMAIN       58     77       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     78     95       2 (POTENTIAL).
FT   DOMAIN       96    105       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    106    127       3 (POTENTIAL).
FT   DOMAIN      128    149       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    150    168       4 (POTENTIAL).
FT   DOMAIN      169    192       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    193    215       5 (POTENTIAL).
FT   DOMAIN      216    241       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    242    265       6 (POTENTIAL).
FT   DOMAIN      266    287       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    288    312       7 (POTENTIAL).
FT   DOMAIN      313    361       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      126    134       INTERACTION WITH G PROTEINS.
FT   DISULFID    104    181       BY SIMILARITY.
SQ   SEQUENCE   361 AA;  41224 MW;  B5A2171F34C9C67B CRC64;

  Query Match             19.0%;  Score 410.5;  DB 1;  Length 361;
  Best Local Similarity   30.2%;  Pred. No. 4.2e-16;
  Matches   99;  Conservative   58;  Mismatches  114;  Indels   57;  Gaps   10;

Qy        106 YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSI-CTTVFYTNLAIADFLFCVT 164
              |   | :  ::|  | |||::|:  | : | ::    : |  ||:: ||| |:| ||
Db         24 YAHHSTARIVMPLHYSLVFIIGLVGNLLALVVIVQNRKKINSTTLYSTNLVISDILFTTA 83

Qy        165 LPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLP 224
              || :|||:  | :|  |: ||| | ::|| | |  :  : |:||:|::|:|||  |  :
Db         84 LPTRIAYYAMGFDWRIGDALCRITALVFYINTYAGVNFMTCLSIDRFIAVVHPLRYNKIK 143

Qy        225 KHTYALVTCGLVWATVFLYMLPFFI---LKQEYYLVQPDITTCHDVHNTCESSS-PFQLY 280
              :  :|   |  ||  ||  ||  |   |||     :  || :  |  |: | |: |
Db        144 RIEHAKGVCIFVWILVFAQTLPLLINPMSKQE-----AERITCMEYPNFEETKSLPWIL- 197

Qy        281 YFISLAFFGFLIPFVLIIYCYAAII-------------------RTLNAYDHRWLWYVK 320
                :   | |:::| ::|: ||: |                    : ||
Db        198 --LGACFIGYVLPLIIILICYSQICCKLFRTAKQNPLTEKSGVNKKALNTI--------- 246

Qy        321 ASLLILVIFTICFAPSNIILIIH------HANYYYNNTDGLYFIYL-IALCLGSLNSCLD 373
              :||:|:| :|| | :: :| |       :|:   :    : | |   :|| : | |:|
Db        247 --ILIIVVFVLCFTPYHVAIIQHMIKKLRFSNFLECSQRHSFQISLHFTVCLMNFNCCMD 304

Qy        374 PFLYFLMSKTRNHSTAYLTKXXNDLREQ 401
              ||:||   |       |  |    |: |
Db        305 PFIYFFACK------GYKRKVMRMLKRQ 326


RESULT 20
```

```
GP17_HUMAN
ID   GP17_HUMAN     STANDARD;      PRT;   367 AA.
AC   Q13304; Q9UDZ6; Q9UE21;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable P2Y purinoceptor GPR17 (P2Y-like receptor) (R12).
GN   GPR17.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=96145150; PubMed=8558062;
RA   Raport C.J., Schweickart V.L., Chantry D., Eddy R.L. Jr., Shows T.B.,
RA   Godiska R., Gray P.W.;
RT   "New members of the chemokine receptor gene family.";
RL   J. Leukoc. Biol. 59:18-23(1996).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Hippocampus;
RX   MEDLINE=98181695; PubMed=9523551;
RA   Blaesius R.H., Weber R.G., Lichter P., Ogilvie A.;
RT   "A novel orphan G protein-coupled receptor primarily expressed in the
RT   brain is localized on human chromosomal band 2q21.";
RL   J. Neurochem. 70:1357-1365(1998).
CC   -!- FUNCTION: Putative receptor for purines coupled to G-proteins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: PRIMARILY EXPRESSED IN BRAIN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U33447; AAB16746.1; -.
DR   EMBL; Y12546; CAA73144.1; -.
DR   EMBL; Z94154; CAB08107.1; -.
DR   EMBL; Z94155; CAB08108.1; -.
DR   HSSP; P34996; 1DDD.
DR   Genew; HGNC:4471; GPR17.
DR   MIM; 603071; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Alternative splicing.
FT   DOMAIN        1     64       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     65     85       1 (POTENTIAL).
FT   DOMAIN       86     92       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     93    113       2 (POTENTIAL).
FT   DOMAIN      114    133       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    134    154       3 (POTENTIAL).
FT   DOMAIN      155    175       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    176    196       4 (POTENTIAL).
FT   DOMAIN      197    223       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    224    244       5 (POTENTIAL).
FT   DOMAIN      245    260       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    261    281       6 (POTENTIAL).
FT   DOMAIN      282    308       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    309    329       7 (POTENTIAL).
FT   DOMAIN      330    367       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    132    209       BY SIMILARITY.
```

```
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    282    282       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     28       MISSING (IN ISOFORM 2).
SQ   SEQUENCE   367 AA;  40989 MW;  132FBE97BE83C60C CRC64;

  Query Match             19.0%;  Score 410.5;  DB 1;  Length 367;
  Best Local Similarity   30.7%;  Pred. No. 4.3e-16;
  Matches  107;  Conservative   48;  Mismatches  132;  Indels   61;  Gaps   12;

Qy         62 APPNSFEEFPPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYL 121
              |||     |  :  |           :| :|                 : |   |  : ||
Db         35 APPGLITNFSLATAE----------QCGQE---------------TPLENMLFASFYL 67

Qy        122 LVFVVGVPANAVTLWMLFFRTRSICT--TVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV 179
              | |:: :  | : || || |      |    ||  :||:||    : || :: || :||:|
Db         68 LDFILALVGNTLALW-LFIRDHKSGTPANVFLMHLAVADLSCVLVLPTRLVYHFSGNHWP 126

Qy        180 FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
              |||: || |  :|| ||| ||  | ||| :|:||||||     | :  || : |  :|
Db        127 FGEIACRLTGFLFYLNMYASIYFLTCISADRFLAIVHPVKSLKLRRPLYAHLACAFLWVV 186

Qy        240 VFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLY-------YFISLAFFGFLI 292
              | : | |  :  |        : | | |   |      |||         :|||   |
Db        187 VAVAMAPLLVSPQ-------TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 231

Qy        293 PFVLIIYCYAAIIRT----LNAYDHRWLWYVKASLLILVIFTICFAPSNI---ILIIHHA 345
              ||:  : ||  |||:     |          |:   ::| || :|| | ::   : ::|:
Db        232 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 291

Qy        346 NYYYNNTDG--LYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
              ::  :      |     |  || |||  ||| :|| :::   |:   |
Db        292 SHGASCATQRILALANRITSCLTSLNGALDPIMYFFVAEKFRHALCNL 339


RESULT 21
P2YR_RAT
ID   P2YR_RAT       STANDARD;      PRT;   373 AA.
AC   P49651;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Insulinoma;
RX   MEDLINE=95298025; PubMed=7779087;
RA   Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;
RT   "Cloning of rat and mouse P2Y purinoceptors.";
RL   Biochem. Biophys. Res. Commun. 211:211-218(1995).
CC   -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC       ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS
CC       OF EXTRACELLULAR ATP ON INSULIN SECRETION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MUSCLE, HEART, LIVER, KIDNEY,
CC       LUNG, BRAIN, SPLEEN, BUT NOT IN TESTIS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U22830; AAA91303.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     52       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    53     74       1 (POTENTIAL).
FT   DOMAIN      75     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    88    109       2 (POTENTIAL).
FT   DOMAIN     110    126       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   127    147       3 (POTENTIAL).
FT   DOMAIN     148    166       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   167    188       4 (POTENTIAL).
FT   DOMAIN     189    218       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   219    238       5 (POTENTIAL).
FT   DOMAIN     239    265       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   266    285       6 (POTENTIAL).
FT   DOMAIN     286    303       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   304    328       7 (POTENTIAL).
FT   DOMAIN     329    373       CYTOPLASMIC (POTENTIAL).
FT   DISULFID   124    202       BY SIMILARITY.
FT   CARBOHYD    11     11       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    27     27       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   373 AA;  42321 MW;  6DDF676287B5E648 CRC64;

  Query Match             19.0%;  Score 409.5;  DB 1;  Length 373;
  Best Local Similarity   28.9%;  Pred. No. 4.9e-16;
  Matches  111;  Conservative   64;  Mismatches  130;  Indels   79;  Gaps   13;

Qy         69 EFPPFSALEG-------------WTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKL 115
              | |:||:                 |  :||                 |:   ::||    |
Db          3 EVPWSAVPNGTDAAFLAGLGSLWGNSTI-----------------ASTAAVSSSFRCAL 44

Qy        116 I---------PAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYTNLAIADFLFCVTL 165
              |          ||:|:|||::|   |:| :||  |  :     :|:  |||:|||| : :||
Db         45 IKTGFQFYYLPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVLTL 104

Qy        166 PFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPK 225
              |  | |: |  :|:||:|:|:    ||: |:| ||| | ||| :||  :|:|    |  |
Db        105 PALIFYYFNKTDWIFGDVMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLK 164

Qy        226 HTYALVTCGLVWATVFLYMLPFFILKQEYY----LVQPDITTCHDVHNTCESSSPFQLYY 281
              |:     |||  | : : |          :|    : :   ||:|      :|  :  |
Db        165 KKNAIYVSVLVWLIVVVAISPIL-----FYSGTGIRKNKTVTCYD-----STSDEYLRSY 214

Qy        282 FI-----SLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKASLLILVIF 329
              ||     ::|   | || |||: ||  |:| |   |         : :: |   :::| :|
Db        215 FIYSMCTTVAM--FCIPLVLILGCYGLIVRALIYKDLDNSPLRRKSIYLV---IIVLTVF 269

Qy        330 TICFAPSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
               : : | :::  ::  |   :          | :|  | :   | |||||:|| ||||
Db        270 AVSYIPFHVMKTMNLRARLDFQTPEMCDFNDRVYATYQVTRGLASLNSCVDPILYFLAGD 329

Qy        383 TRNHSTAYLTKXXNDLREQGQPSQ 406
              |     :  |:  :   |    |:
Db        330 TFRRRLSRATRKASRRSEANLQSK 353


RESULT 22
P2Y5_CHICK
ID   P2Y5_CHICK     STANDARD;      PRT;   308 AA.
AC   P32250;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   P2Y purinoceptor 5 (P2Y5) (Purinergic receptor 5) (6H1).
GN   P2RY5.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
```

```
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=T-cell;
RX   MEDLINE=93329058; PubMed=8393036;
RA   Kaplan M.H., Smith D.I., Sundick R.S.;
RT   "Identification of a G protein coupled receptor induced in activated
RT   T cells.";
RL   J. Immunol. 151:628-636(1993).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: INDUCED IN ACTIVATED T-CELLS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L06109; AAB06587.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
KW   Palmitate.
FT   DOMAIN        1     16       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     17     43       1 (POTENTIAL).
FT   DOMAIN       44     52       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     53     76       2 (POTENTIAL).
FT   DOMAIN       77     89       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     90    109       3 (POTENTIAL).
FT   DOMAIN      110    130       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    131    151       4 (POTENTIAL).
FT   DOMAIN      152    178       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    179    206       5 (POTENTIAL).
FT   DOMAIN      207    224       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    225    250       6 (POTENTIAL).
FT   DOMAIN      251    269       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    270    289       7 (POTENTIAL).
FT   DOMAIN      290    308       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   LIPID       281    281       PALMITATE (BY SIMILARITY).
FT   DISULFID     86    165       POTENTIAL.
SQ   SEQUENCE   308 AA;  35597 MW;  4214E969633B6F7D CRC64;

  Query Match             18.8%;  Score 406.5;  DB 1;  Length 308;
  Best Local Similarity   33.5%;  Pred. No. 6.1e-16;
  Matches  108;  Conservative   52;  Mismatches  127;  Indels   35;  Gaps   11;

Qy        107 LTSSLSTK------LIPAIYLLVFVVGVPANAVTLWMLFF--RTRSICTTVFYTNLAIAD 158
              ::|: ||:     |   :: :|||:|: || | ::::  |  : |:  || :  ||||:|
Db          2 VSSNCSTEDSFKYTLYGCVFSMVFVLGLIANCVAIYIFTFTLKVRNE-TTTYMLNLAISD 60

Qy        159 FLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF 218
               ||  ||||:| |:    || ||:|||: :  :|| ||| ||| | |||::|:|||||||
Db         61 LLFVFTLPFRI-YYFVVRNWPFGDVLCKISVTLFYTNMYGSILFLTCISVDRFLAIVHPF 119

Qy        219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESS-SPF 277
                : |     | : |  || ||     |    :        :  || :  |  ||:   :
Db        120 RSKTLRTKRNARIVCVAVWITVLAGSTPASFFQSTNRQNNTEQRTCFE--NFPESTWKTY 177

Qy        278 QLYYFISLAFFGFLIPFVLIIYCYAAIIRTLN---AYDHRWLWYVKASLLI---LVIFTI 331
                  |  :   || || :| : |   ::||||         |   |   :|   ||||
Db        178 LSRIVIFIEIVGFFIPLILNVTCSTMVLRTLNKPLTLSRNKLSKKKVLKMIFVHLVIFCF 237

Qy        332 CFAPSNIILIIH---HANYYYNNT--DGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNH 386
              || | || ||::      : | :   :  :| : ||:   | | || :|:  | |
Db        238 CFVPYNITLILYSLMRTQTWINCSVVTAVRTMYPVTLCIAVSNCCFDPIVYYFTSDTN-- 295

Qy        387 STAYLTKXXNDLREQGQPSQRT 408
                      ::| :: |  | | |
Db        296 ---------SELDKKQQVHQNT 308


RESULT 23
P2YR_HUMAN
ID   P2YR_HUMAN     STANDARD;      PRT;   373 AA.
AC   P47900;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=96257237; PubMed=8666290;
RA   Leon C., Vial C., Cazenave J.-P., Gachet C.;
RT   "Cloning and sequencing of a human cDNA encoding endothelial P2Y1
RT   purinoceptor.";
RL   Gene 171:295-297(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96158962; PubMed=8579591;
RA   Ayyanathan K., Tania W., Harbansjit S., Raghbir A.S., Barnard E.A.,
RA   Kunapuli S.P.;
RT   "Cloning and chromosomal localization of the human P2Y1
RT   purinoceptor.";
RL   Biochem. Biophys. Res. Commun. 218:783-788(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96205320; PubMed=8630005;
RA   Janssens R., Communi D., Pirotton S., Samson M., Parmentier M.,
RA   Boeynaems J.M.;
RT   "Cloning and tissue distribution of the human P2Y1 receptor.";
RL   Biochem. Biophys. Res. Commun. 221:588-593(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RA   Leon C., Vial C., Weber J., Cazenave J.-P., Gacher C.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 95-373 FROM N.A., FUNCTION, AND INHIBITION STUDIES.
RC   TISSUE=Platelet;
RX   MEDLINE=98113162; PubMed=9442040;
RA   Jin J., Daniel J.L., Kunapuli S.P.;
RT   "Molecular basis for ADP-induced platelet activation. II. The P2Y1
RT   receptor mediates ADP-induced intracellular calcium mobilization and
RT   shape change in platelets.";
RL   J. Biol. Chem. 273:2030-2034(1998).
CC   -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC       ATP AND ADP. IN PLATELETS BINDING TO ADP LEADS TO MOBILIZATION OF
CC       INTRACELLULAR CALCIUM IONS VIA ACTIVATION OF PHOSPHOLIPASE C, A
CC       CHANGE IN PLATELET SHAPE, AND PROBABLY TO PLATELET AGGREGATION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- INDUCTION: REPRESSED BY THE P2Y1 RECEPTOR-SPECIFIC ANTAGONISTS
CC       A3P5PS, A3P5P AND A2P5P. THESE INHIBIT CALCIUM ION MOBILIZATION
CC       AND SHAPE CHANGE IN PLATELETS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; Z49205; CAA89066.1; -.
DR   EMBL; U42030; AAA97873.1; -.
DR   EMBL; U42029; AAA97872.1; -.
DR   EMBL; S81950; AAB47091.1; -.
DR   EMBL; AJ006945; CAA07339.1; -.
DR   EMBL; AF018284; AAB94556.1; -.
DR   HSSP; P34996; 1DDD.
DR   Genew; HGNC:8539; P2RY1.
DR   MIM; 601167; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Platelet;
KW   Blood coagulation.
FT   DOMAIN        1     52       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     53     74       1 (POTENTIAL).
FT   DOMAIN       75     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    109       2 (POTENTIAL).
FT   DOMAIN      110    126       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    127    147       3 (POTENTIAL).
FT   DOMAIN      148    166       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    167    188       4 (POTENTIAL).
FT   DOMAIN      189    218       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    219    238       5 (POTENTIAL).
FT   DOMAIN      239    265       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    266    285       6 (POTENTIAL).
FT   DOMAIN      286    303       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    304    328       7 (POTENTIAL).
FT   DOMAIN      329    373       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    124    202       BY SIMILARITY.
FT   CARBOHYD     11     11       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     27     27       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    138    138       MISSING (IN REF. 1).
SQ   SEQUENCE   373 AA;  42071 MW;  4DC7C668B4145392 CRC64;

  Query Match             18.5%;  Score 399;  DB 1;  Length 373;
  Best Local Similarity   29.7%;  Pred. No. 1.8e-15;
  Matches  113;  Conservative   66;  Mismatches  143;  Indels   58;  Gaps   14;

Qy         52 PTLPIKT---FRGAPPNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLT 108
              | :|  |    |   | :|          |  :|:  .          ||   | ::    ||
Db          7 PAVPNGTDAAFLAGPGSS----------WGNSTV----------ASTAAVSSSFKCALT 45

Qy        109 -SSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYTNLAIADFLFCVTLP 166
               :      :||:|:|||::|   |:| :||  |  :     :|:  |||:|||| : |||
Db         46 KTGFQFYYLPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVLTLP 105

Qy        167 FKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKH 226
                | |: |  :|:||: :|:    ||: |:| ||| | ||| :||  :|:|     |  |
Db        106 ALIFYYFNKTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKK 165

Qy        227 TYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDIT-TCHDVHNTCESSSPFQLYYFI-- 283
                |:    |||  | : : |  ||       |: : | ||:|      :|  :   |||
Db        166 KNAICISVLVWLIVVVAISP--ILFYSGTGVRKNKTITCYDT-----TSDEYLRSYFIYS 218

Qy        284 ---SLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKASLLILVIFTICF 333
                 ::|   | :| |||: ||  |:| |   |        : :: |   :::| :| : :
Db        219 MCTTVAM--FCVPLVLILGCYGLIVRALIYKDLDNSPLRRKSIYLV---IIVLTVFAVSY 273

Qy        334 APSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNH 386
                | :::  ::  |   :          | :|  | :   | |||||:|| ||||   |
Db        274 IPFHVMKTMNLRARLDFQTPAMCAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRR 333

Qy        387 STAYLTKXXNDLREQGQPSQ 406
               :  |:  :   |    |:
Db        334 RLSRATRKASRRSEANLQSK 353
```

```
RESULT 24
P2YR_BOVIN
ID   P2YR_BOVIN     STANDARD;      PRT;   373 AA.
AC   P48042;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aortic endothelium;
RX   MEDLINE=95352058; PubMed=7626079;
RA   Henderson D.J., Elliot D.G., Smith G.M., Webb T.E., Dainty I.A.;
RT   "Cloning and characterisation of a bovine P2Y receptor.";
RL   Biochem. Biophys. Res. Commun. 212:648-656(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus callosum;
RX   MEDLINE=99064562; PubMed=9848096;
RA   Deng G., Matute C., Kumar C.K., Fogarty D.J., Miledi R.;
RT   "Cloning and expression of a P2y purinoceptor from the adult bovine
RT   corpus callosum.";
RL   Neurobiol. Dis. 5:259-270(1998).
CC   -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC       ATP AND ADP.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X87628; CAA60958.1; -.
DR   EMBL; U34041; AAC78275.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     52       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     53     74       1 (POTENTIAL).
FT   DOMAIN       75     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    109       2 (POTENTIAL).
FT   DOMAIN      110    126       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    127    147       3 (POTENTIAL).
FT   DOMAIN      148    166       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    167    188       4 (POTENTIAL).
FT   DOMAIN      189    218       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    219    238       5 (POTENTIAL).
FT   DOMAIN      239    265       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    266    285       6 (POTENTIAL).
FT   DOMAIN      286    303       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    304    328       7 (POTENTIAL).
FT   DOMAIN      329    373       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    124    202       BY SIMILARITY.
FT   CARBOHYD     11     11       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     27     27       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   373 AA;  42287 MW;  9270A7175C0BDA76 CRC64;
```

```
  Query Match           18.5%;  Score 398;  DB 1;  Length 373;
  Best Local Similarity   28.5%;  Pred. No. 2e-15;
  Matches  111;  Conservative   66;  Mismatches  136;  Indels   76;  Gaps   15;

Qy         45 DTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEGWTGATI--TVKIKCPEESASHLHVKNA 102
              ||  || |        |:|              |  :|:  |  :   | : |
Db         14 DTAFLADP--------GSP--------------WGNSTVTSTAAVASPFKCA-------- 43

Qy        103 TMGYLT-SSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYTNLAIADFL 160
                 ||  :      :||:|:|||::|   |:| :||  |  :     :|:. |||:||||
Db         44 ----LTKTGFQFYYLPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFL 99

Qy        161 FCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTY 220
              : :|||  | |: |  :|:||: :|:     ||: |:| ||| | ||| :||  :|:|
Db        100 YVLTLPALIFYYFNKTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKS 159

Qy        221 RGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY----LVQPDITTCHDVHNTCESSSP 276
               |  |   |:    |||  | : : |·      :|    : :    ||:|       :|
Db        160 LGRLKKKNAVYISVLVWLIVVVGISPIL-----FYSGTGIRKNKTITCYDT-----TSDE 209

Qy        277 FQLYYFI-----SLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKASLL 324
              :  |||      ::|   | :| |||: ||  |:| |   |         : :: |   ::
Db        210 YLRSYFIYSMCTTVAM--FCVPLVLILGCYGLIVRALIYKDLDNSPLRRKSIYLV---II 264

Qy        325 ILVIFTICFAPSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDPFLY 377
              :| :|  : : | :::  ::  |   :          | :|  | :   | |||||:|| ||
Db        265 VLTVFAVSYIPFHVMKTMNLRARLDFQTPEMCAFNDRVYATYQVTRGLASLNSCVDPILY 324

Qy        378 FLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
              ||   |      :  |:  :   |    |:
Db        325 FLAGDTFRRRLLSRATRKASRRSEANLQSK 353
```

```
RESULT 25
P2YR_MOUSE
ID   P2YR_MOUSE     STANDARD;      PRT;   373 AA.
AC   P49650;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Insulinoma;
RX   MEDLINE=95298025; PubMed=7779087;
RA   Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;
RT   "Cloning of rat and mouse P2Y purinoceptors.";
RL   Biochem. Biophys. Res. Commun. 211:211-218(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RA   Leon C.;
RT   "Thromboresistance in P2Y1 receptor knockout mice.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC       ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS
CC       OF EXTRACELLULAR ATP ON INSULIN SECRETION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U22829; AAA91302.1; -.
DR   EMBL; AJ245636; CAB57317.1; -.
DR   HSSP; P34996; 1DDD.
DR   MGD; MGI:105049; P2ry1.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     52       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     53     74       1 (POTENTIAL).
FT   DOMAIN       75     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    109       2 (POTENTIAL).
FT   DOMAIN      110    126       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    127    147       3 (POTENTIAL).
FT   DOMAIN      148    166       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    167    188       4 (POTENTIAL).
FT   DOMAIN      189    218       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    219    238       5 (POTENTIAL).
FT   DOMAIN      239    265       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    266    285       6 (POTENTIAL).
FT   DOMAIN      286    303       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    304    328       7 (POTENTIAL).
FT   DOMAIN      329    373       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    124    202       BY SIMILARITY.
FT   CARBOHYD     11     11       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     27     27       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   373 AA;  42212 MW;  944125E9F4560BB3 CRC64;
```

```
  Query Match           18.4%;  Score 397.5;  DB 1;  Length 373;
  Best Local Similarity   31.1%;  Pred. No. 2.2e-15;
  Matches   98;  Conservative   61;  Mismatches  117;  Indels   39;  Gaps   11;

Qy        116 IPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYTNLAIADFLFCVTLPFKIAYHLN 174
              :||:|:|||::|   |:| :||  |  :     :|:  |||:|||:: :|||  | |: |
Db         54 LPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVLTLPALIFYYFN 113

Qy        175 GNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCG 234
                :|:||: :|:    ||: |:| ||| | ||| :||  :|:|    | ·|   |:
Db        114 KTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKKKNAIYVSV 173

Qy        235 LVWATVFLYMLPFFILKQEYY----LVQPDITTCHDVHNTCESSSPFQLYYFI-----SL 285
              |||  | : : |       :|       :    ||:|      :|: :  |||     ::
Db        174 LVWLIVVVAISPIL-----FYSGTGTRKNKTVTCYDT-----TSNDYLRSYFIYSMCTTV 223

Qy        286 AFFGFLIPFVLIIYCYAAIIRTL--NAYDH-----RWLWYVKASLLILVIFTICFAPSNI 338
              |   | || |||: || ·|:: |  |  |  :     : :: |   :::| :| : : | ::
Db        224 AM--FCIPLVLILGCYGLIVKALIYNDLDNSPLRRKSIYLV---IIVLTVFAVSYIPFHV 278

Qy        339 ILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
              :  ::  |   :          | :|  | :  | |||||:|| ||||   |     :
Db        279 MKTMNLRARLDFQTPEMCDFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRRLSRA 338

Qy        392 TKXXNDLREQGQPSQ 406
              |:  :   |    |:
Db        339 TRKASRRSEANLQSK 353
```

```
RESULT 26
P2Y9_HUMAN
ID   P2Y9_HUMAN     STANDARD;      PRT;   370 AA.
AC   Q99677; O15132;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   P2Y purinoceptor 9 (P2Y9) (Purinergic receptor 9) (G protein-coupled
DE   receptor GPR23) (P2Y5-like receptor).
GN   GPR23 OR P2RY9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97225799; PubMed=9073069;
RA   O'Dowd B.F., Nguyen T., Jung B.P., Marchese A., Cheng R.,
RA   Heng H.H.Q., Kolakowski L.F. Jr., Lynch K.R., George S.R.;
RT   "Cloning and chromosomal mapping of four putative novel human
RT   G-protein-coupled receptor genes.";
RL   Gene 187:75-81(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Bohm S.K., Khitin L.M., Payan D.P., Bunnett N.W.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97366605; PubMed=9223435;
RA   Janssens R., Boeynaems J.M., Godart M., Communi D.;
RT   "Cloning of a human heptahelical receptor closely related to the P2Y5
RT   receptor.";
RL   Biochem. Biophys. Res. Commun. 236:106-112(1997).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: NOT DETECTED IN THE BRAIN REGIONS THALAMUS,
CC       PUTAMEN, CAUDATE, FRONTAL CORTEX, PONS, HYPOTHALAMUS, HIPPOCAMPUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U66578; AAC51301.1; -.
DR   EMBL; U90323; AAB62087.1; -.
DR   EMBL; U90322; AAB62088.1; -.
DR   EMBL; AF005419; AAB66322.1; -.
DR   HSSP; P34996; 1DDD.
DR   Genew; HGNC:4478; GPR23.
DR   MIM; 300086; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     43       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     44     64       1 (POTENTIAL).
FT   DOMAIN       65     73       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     74     94       2 (POTENTIAL).
FT   DOMAIN       95    112       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    113    133       3 (POTENTIAL).
FT   DOMAIN      134    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    176       4 (POTENTIAL).
FT   DOMAIN      177    203       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    204    224       5 (POTENTIAL).
FT   DOMAIN      225    254       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    255    275       6 (POTENTIAL).
FT   DOMAIN      276    294       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    295    315       7 (POTENTIAL).
FT   DOMAIN      316    370       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    111    188       BY SIMILARITY.
FT   CARBOHYD     15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     28     28       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    183    183       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    192    192       F -> L (IN REF. 3).
SQ   SEQUENCE   370 AA;  41895 MW;  20857F52A3929E48 CRC64;

  Query Match          17.8%;  Score 384.5;  DB 1;  Length 370;
  Best Local Similarity  30.4%;  Pred. No. 1.1e-14;
  Matches   95;  Conservative   60;  Mismatches  121;  Indels   37;  Gaps   10;

Qy         91 EESASHLHVK--NATMG---YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSI 145
              ::| | |  :  |||      :  |    |  |:| :||::|:  |:|:|::  || :
Db         12 QDSNSSLRPRLGNATANNTCIVDDSFKYNLNGAVYSVVFILGLITNSVSLFVFCFRMKMR 71

Qy        146 C-TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLA 204
               | :| ||||::| ||  |||||| |: |  :| ||: ||: :   |  |:| |:| |
Db         72 SETAIFITNLAVSDLLFVCTLPFKIFYNFN-RHWPFGDTLCKISGTAFLTNIYGSMLFLT 130

Qy        205 CISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTC 264
              |||::|:||||:||  | :     : : |  ||  |    :   :              |
Db        131 CISVDRFLAIVYPFRSRTIRTRRNSAIVCAGVWILVLSGGISASLF---------STTNV 181

Qy        265 HDVHNTCESSSPFQLY--YFISLAFF----GFLIPFVLIIYCYAAIIR------TLNAYD 312
              ::   ||       :::  |   :  |    ||:|| :| : | : ::|      ||:
Db        182 NNATTTCFEGFSKRVWKTYLSKITIFIEVVGFIIPLILNVSCSSVVLRTLRKPATLSQIG 241

Qy        313 HRWLWYVKASLLILVIFTICFAPSNIILIIHH-------ANYYYNNTDGLYFIYLIALCL 365
                    :|   : : :| :|| | | :|  ::          | :      :  :| | |||
Db        242 TNKKKVLKMITVHMAVFVVCFVPYNSVLFLYALVRSQAITNCFLERFAKI--MYPITLCL 299

Qy        366 GSLNSCLDPFLYF 378
               :|| | |||:|:
Db        300 ATLNCCFDPFIYY 312


RESULT 27
P2Y4_HUMAN
ID   P2Y4_HUMAN     STANDARD;      PRT;   365 AA.
AC   P51582;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   P2Y purinoceptor 4 (P2Y4) (Uridine nucleotide receptor) (UNR) (P2P).
GN   P2RY4 OR NRU.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96125055; PubMed=8537336;
RA   Communi D., Pirotton S., Parmentier M., Boeynaems J.-M.;
RT   "Cloning and functional expression of a human uridine nucleotide
RT   receptor.";
RL   J. Biol. Chem. 270:30849-30852(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96125054; PubMed=8537335;
RA   Nguyen T., Erb L., Weisman G.A., Marchese A., Heng H.H.Q.,
RA   Garrad R.C., George S.R., Turner J.T., O'Dowd B.F.;
RT   "Cloning, expression, and chromosomal localization of the human
RT   uridine nucleotide receptor gene.";
RL   J. Biol. Chem. 270:30845-30848(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=96197801; PubMed=8617367;
RA   Stam N.J., Klomp J., van der Heuvel M., Olijve W.;
RT   "Molecular cloning and characterization of a novel orphan receptor
RT   (P2P) expressed in human pancreas that shows high structural homology
RT   to the P2U purinoceptor.";
RL   FEBS Lett. 384:260-264(1996).
CC   -!- FUNCTION: Receptor for UTP and UDP coupled to G-proteins that
CC       activate a phosphatidylinositol-calcium second messenger system.
CC       Not activated by ATP or ADP.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Pancreas.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X91852; CAA62963.1; -.
DR   EMBL; U40223; AAC50347.1; -.
DR   EMBL; X96597; CAA65415.1; -.
DR   HSSP; P34996; 1DDD.
DR   Genew; HGNC:8542; P2RY4.
DR   MIM; 300038; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Polymorphism.
FT   DOMAIN       1     34       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     35    61       1 (POTENTIAL).
FT   DOMAIN       62    72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73    95       2 (POTENTIAL).
FT   DOMAIN       96    112      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     113   131      3 (POTENTIAL).
FT   DOMAIN       132   154      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     155   174      4 (POTENTIAL).
FT   DOMAIN       175   196      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     197   222      5 (POTENTIAL).
FT   DOMAIN       223   246      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     247   269      6 (POTENTIAL).
FT   DOMAIN       270   287      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     288   309      7 (POTENTIAL).
FT   DOMAIN       310   365      CYTOPLASMIC (POTENTIAL).
FT   DISULFID     108   185      BY SIMILARITY.
FT   VARIANT      168   168      V -> M (IN DBSNP:1152186).
FT                               /FTId=VAR_011854.
FT   VARIANT      178   178      N -> T (IN DBSNP:1152187).
FT                               /FTId=VAR_011855.
FT   VARIANT      191   191      P -> L (IN DBSNP:1152188).
FT                               /FTId=VAR_011856.
FT   CONFLICT      86    86      L -> V (IN REF. 2).
FT   CONFLICT     234   234      S -> A (IN REF. 2).
SQ   SEQUENCE   365 AA;  40963 MW;  23E0AFED3B7BDEED CRC64;

  Query Match             17.7%;  Score 382;  DB 1;  Length 365;
  Best Local Similarity   31.8%;  Pred. No. 1.5e-14;
  Matches  100;  Conservative   47;  Mismatches  127;  Indels   40;  Gaps   10;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFYT-NLAIADFLFCVTLPFKIAYHL 173
              |:|  | :|||:|:  || |||:  || |    |  |  :||::| |: ::||  | |:
Db         37 LLPVSYAVVFVLGLGLNAPTLWLFIFRLRPWDATATYMFHLALSDTLYVLSLPTLIYYYA 96

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
                |:| ||  :|:    :|| |:|||:| | |||::||| | ||       :   | : |
Db         97 AHNHWPFGTEICKFVRFLFYWNLYCSVLFLTCISVHRYLGICHPLRALRWGRPRLAGLLC 156

Qy        234 GLVWATVFLYMLP--FFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFL 291
                ||  |   ::|  ||:        :     |||      |    | |   | | | ||
Db        157 LAVWLVVAGCLVPNLFFVTTSN----KGTTVLCHDTTRPEE----FDHYVHFSSAVMGLL 208

Qy        292 --IPFVLIIYCYAAIIRTL------NAYDHRWLWYVKASLLILVIFTICFAPSNIILIIH 343
                :| :: : ||  : | |      :|    |  ::   ::| :| :|| | :|   |:
Db        209 FGVPCLVTLVCYGLMARRLYQPLPGSAQSSSRLRSLRTIAVVLTVFAVCFVPFHITRTIY 268

Qy        344 H------ANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXND 397
                :      |:    |   :  :| :   | | |||||| || |             |
Db        269 YLARLLEADCRVLNI--VNVVYKVTRPLASANSCLDPVLYLLTG----------DKYRRQ 316

Qy        398 LRE---QGQPSQRT 408
                ||:    |:|  ||
Db        317 LRQLCGGGKPQPRT 330


RESULT 28
CLT2_RAT
ID   CLT2_RAT       STANDARD;      PRT;   309 AA.
AC   Q924T9;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 2 (CysLTR2) (RSBPT32).
GN   CYSLTR2 OR CYSLT2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kamohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
RA   Ohishi T., Soga T., Matsushime H., Furuichi K.;
RT   "Characterization of the cloned rat and porcine cysteinyl leukotriene
RT   receptors.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC       mediated via a G-protein that activates a phosphatidylinositol-
CC       calcium second messenger system (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB052661; BAB60816.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27    47       1 (POTENTIAL).
FT   DOMAIN       48    56       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     57    77       2 (POTENTIAL).
FT   DOMAIN       78    98       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     99    119      3 (POTENTIAL).
FT   DOMAIN       120   137      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     138   158      4 (POTENTIAL).
FT   DOMAIN       159   187      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     188   208      5 (POTENTIAL).
FT   DOMAIN       209   229      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     230   250      6 (POTENTIAL).
FT   DOMAIN       251   271      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     272   292      7 (POTENTIAL).
FT   DOMAIN       293   309      CYTOPLASMIC (POTENTIAL).
FT   DISULFID      95   171      BY SIMILARITY.
FT   CARBOHYD      14    14      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     166   166      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     167   167      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   309 AA;  35507 MW;  A122AC8177879D56 CRC64;

  Query Match             17.6%;  Score 380.5;  DB 1;  Length 309;
  Best Local Similarity   31.4%;  Pred. No. 1.6e-14;
  Matches   96;  Conservative   55;  Mismatches  110;  Indels   45;  Gaps   11;

Qy         95 SHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRT--RSICTTVFYT 152
              |:   || |:      :      | |||::|| |    |  ::: :| :|  :|    ||
Db          8 SYYSDKNCTI----ENFKRDFYPIIYLIIFVWGALGNGFSIY-VFLQTYKKSTSVNVFML 62

Qy        153 NLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYL 212
              ||||:|||| ||||:  |:  |::|:||:  ||  :   | ||| ||  |  :|| |:|
Db         63 NLAISDFLFISTLPFRADYNFRGSDWIFGDWACRIMSYSLYVNMYTSIYFLTVLSIVRFL 122

Qy        213 AIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILK--QEYYLVQPDITTCHDVHNT 270
              | |||   :     | : ||::|  ||:      :||  ||    : : | | :::
```

```
Db        123 ATAHPFQMLHITSVRSAWILCGIIW--VFIMASSGLLLKHGQE---KKNNTTLCFELN-- 175

Qy        271 CESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWYVKASL------- 323
              :     :  :|:|   |||:|| ::  ||  ||| |          :|  :
Db        176 LQKFKNLVILNYIALG-VGFLLPFFILTICYLLIIRVL----------LKVEIPESGPRD 224

Qy        324 ----------LILVIFTICFAPSNIILIIHHANYYYNN-TDGLYFIYLIALCLGSLNSCL 372
                       : ::|| :|| | : :  ||   :  ::  | |:   :| | | : |||
Db        225 AQRKALTTIVIAMIIFLLCFLPYHALRTIHLVTWDADSCMDELHKATVITLTLAAANSCF 284

Qy        373 DPFLYF 378
                 :||||:
Db        285 NPFLYY 290


RESULT 29
CLT2_PIG
ID   CLT2_PIG       STANDARD;      PRT;   345 AA.
AC   Q95N03;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 2 (CysLTR2).
GN   CYSLTR2 OR CYSLT2.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kamohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
RA   Ohishi T., Soga T., Matsushime H., Furuichi K.;
RT   "Characterization of the cloned rat and porcine cysteinyl leukotriene
RT   receptors.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC       mediated via a G-protein that activates a phosphatidylinositol-
CC       calcium second messenger system (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB052662; BAB60817.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     43       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     44     64       1 (POTENTIAL).
FT   DOMAIN       65     73       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     74     94       2 (POTENTIAL).
FT   DOMAIN       95    124       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    125    145       3 (POTENTIAL).
FT   DOMAIN      146    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    175       4 (POTENTIAL).
FT   DOMAIN      176    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206    226       5 (POTENTIAL).
FT   DOMAIN      227    246       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    247    267       6 (POTENTIAL).
FT   DOMAIN      268    287       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    288    308       7 (POTENTIAL).
FT   DOMAIN      309    345       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    112    188       BY SIMILARITY.
FT   CARBOHYD     20     20       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     29     29       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    184    184       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   345 AA;  39410 MW;  5D1B1FB89BB95905 CRC64;

  Query Match             17.6%;  Score 379.5;  DB 1;  Length 345;
  Best Local Similarity   28.7%;  Pred. No. 1.9e-14;
  Matches   97;  Conservative   68;  Mismatches  126;  Indels   47;  Gaps   13;

Qy         92 ESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFR--TRSICTTV 149
              |  | |    |:      | :   :  | :||::|: |    |  ::: :| :   :|     |
Db         18 EPNSTLGNHNSNRSCTTENFKREFYPIVYLVIFIWGALGNGFSIY-VFLKPYKKSTSVNV 76

Qy        150 FYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISIN 209
              |  ||||:| || :||||:: |:| |:| :||:  ||  :   | ||| ||  |  :|:
Db         77 FMLNLAISDLLFTITLPFRVDYYLRGSNXIFGDTPCRIMSYSMYVNMYSSIYFLTVLSVV 136

Qy        210 RYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLV------QPDITT 263
              |:|| ||||          | : ||::|          ||:     |:      : ::|
Db        137 RFLATVHPFRLLHTTSIKNAWILCGVIW---------IFIMASSTVLLKNGSEQKDNVTL 187

Qy        264 CHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYD---------HR 314
              | ::::     :  :   :::|   ||::||  :  ||  ||| |   :           ||
Db        188 CLELNS--NKVTKLKTMNYVALV-VGFVLPFGTLSICYLLIIRALLKVEVPESGLRLSHR 244

Qy        315 WLWYVKASLLI---LVIFTICFAPSNIILIIHHANYYYNN-TDGLYFIYLIALCLGSLNS 370
                  || : :   |:|| :|| | ::: :|   :  :   | |:    : | | : ||
Db        245 -----KALITVIIALIIFLLCFLPYHVLRTLHLLEWKADKCKDRLHKAVAVTLALAAANS 299

Qy        371 CLDPFLYFLMSKTRNHSTAYLTKXXNDLREQGQPSQRT 408
              | :||||:   :       :  :  : || :|:| |:|
Db        300 CFNPFLYYFAGEN------FKDRLKSALR-KGRP-QKT 329


RESULT 30
P2Y4_RAT
ID   P2Y4_RAT       STANDARD;      PRT;   361 AA.
AC   O35811;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   P2Y purinoceptor 4 (P2Y4).
GN   P2RY4 OR P2Y4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RA   Bogdanov Y.D., Wildman S., King B.F., Burntock G.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=98421785; PubMed=9751165;
RA   Webb T.E., Henderson D., Roberts J.A., Barnard E.A.;
RT   "Molecular cloning and characterization of the rat P2Y4 receptor.";
RL   J. Neurochem. 71:1424-1434(1998).
CC   -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
CC       activate a phosphatidylinositol-calcium second messenger system.
CC       Not activated by ADP or UDP.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Widely expressed at low levels. In brain,
CC       higher expression in the pineal gland and ventricular system.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; Y14705; CAA75007.1; -.
DR   EMBL; Y11433; CAA72241.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     30       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     31    58       1 (POTENTIAL).
FT   DOMAIN       59    68       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     69    91       2 (POTENTIAL).
FT   DOMAIN       92    108      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     109   127      3 (POTENTIAL).
FT   DOMAIN       128   149      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     150   170      4 (POTENTIAL).
FT   DOMAIN       171   192      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     193   218      5 (POTENTIAL).
FT   DOMAIN       219   242      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     243   265      6 (POTENTIAL).
FT   DOMAIN       266   283      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     284   305      7 (POTENTIAL).
FT   DOMAIN       306   361      CYTOPLASMIC (POTENTIAL).
FT   DISULFID     104   181      BY SIMILARITY.
FT   CARBOHYD     175   175      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   361 AA;  40893 MW;  0377F96E54B449A3 CRC64;

  Query Match             17.5%;  Score 378;  DB 1;  Length 361;
  Best Local Similarity   31.4%;  Pred. No. 2.4e-14;
  Matches   98;  Conservative   51;  Mismatches  127;  Indels   36;  Gaps   10;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFYT-NLAIADFLFCVTLPFKIAYHL 173
              |:|  | :|||:|:  || ||| :  || |    |  |  :||::| |: ::||  : |:
Db         33 LLPMSYAVVFVLGLALNAPTLWLFLFRLRPWDATATYMFHLALSDTLYVLSLPTLVYYYA 92

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
              |:| ||  ||:    :|| |:|||:| | |||::||| | ||       :  :| : |
Db         93 ARNHWPFGTGLCKFVRFLFYWNLYCSVLFLTCISVHRYLGICHPLRAIRWGRPRFASLLC 152

Qy        234 GLVWATVFLYMLP--FFILKQEYYLVQPDITT--CHDVHNTCESSSPFQLYYFISLAFFG 289
               ||  |   ::|  ||:          : ||  |||   |        :|:  ::
Db        153 LGVWLVVAGCLVPNLFFV------TTNANGTTILCHD--TTLPEEFDHYVYFSSAVMVLL 204

Qy        290 FLIPFVLIIYCYAAIIRTL------NAYDHRWLWYVKASLLILVIFTICFAPSNIILIIH 343
              | :||:: : ||  : | |              |  ::   ::| :| :|| | :|   |:
Db        205 FGLPFLITLVCYGLMARRLYRPLPGAGQSSSRLRSLRTIAVVLTVFAVCFVPFHITRTIY 264

Qy        344 H------ANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNHSTAYLTKXXN 396
              :      |: :  |   :  :| :   | | |||||| || |   | ||
Db        265 YQARLLQADCHVLNI--VNVVYKVTRPLASANSCLDPVLYLFTGDKYRNQ--------LQ 314

Qy        397 DLREQGQPSQRT 408
               |   :|  ||
Db        315 QLCRGSKPKPRT 326


RESULT 31
CLT2_HUMAN
ID   CLT2_HUMAN     STANDARD;      PRT;   346 AA.
AC   Q9NS75; Q9HCQ2;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 2 (CysLTR2) (PSEC0146) (HG57) (HPN321).
GN   CYSLTR2 OR CYSLT2 OR CYSLT2R.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=20374466; PubMed=10913337;
RA   Takasaki J., Kamohara M., Matsumoto M., Saito T., Sugimoto T., Ota T.,
RA   Nishikawa T., Kawai Y., Masuho Y., Isogai T., Suzuki Y., Sugano S.,
RA   Furuichi K.;
RT   "The molecular characterization and tissue distribution of the human
RT   cysteinyl leukotriene CysLT2 receptor.";
RL   Biochem. Biophys. Res. Commun. 274:316-322(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20459128; PubMed=10851239;
RA   Heise C.E., O'Dowd B.F., Figueroa D.J., Sawyer N., Nguyen T.,
RA   Im D.-S., Stocco R., Bellefeuille J.N., Abramovitz M., Cheng R.,
RA   Williams D.L. Jr., Zeng Z., Liu Q., Ma L., Clements M.K., Coulombe N.,
RA   Liu Y., Austin C.P., George S.R., O'Neill G.P., Metters K.M.,
RA   Lynch K.R., Evans J.F.;
RT   "Characterization of the human cysteinyl leukotriene 2 receptor.";
RL   J. Biol. Chem. 275:30531-30536(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20545741; PubMed=11093801;
RA   Nothacker H.-P., Wang Z., Zhu Y., Reinscheid R.K., Lin S.H.S.,
RA   Civelli O.;
RT   "Molecular cloning and characterization of a second human cysteinyl
RT   leukotriene receptor: discovery of a subtype selective agonist.";
RL   Mol. Pharmacol. 58:1601-1608(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Dunn M.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 17-346 FROM N.A.
RA   Suga H.;
RT   "Homo sapiens cysteinyl leukotriene receptor 1 like receptor.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC       mediated via a G-protein that activates a phosphatidylinositol-
CC       calcium second messenger system. Stimulation by BAY u9773, a
CC       partial agonist, induces specific contractions of pulmonary veins
CC       and might also have an indirect role in the relaxation of the
CC       pulmonary vascular endothelium. The rank order of affinities for
CC       the leukotrienes is LTC4 = LTD4 >> LTE4.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in the
CC       heart, placenta, spleen, peripheral blood leukocytes and adrenal
CC       gland. In lung, expressed in the interstitial macrophages, and
CC       slightly in smooth muscle cells.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB038269; BAB03601.1; -.
DR   EMBL; AF254664; AAG17281.1; -.
DR   EMBL; AF279611; AAK69485.1; -.
DR   EMBL; AL137118; CAC29102.1; -.
DR   EMBL; AB041644; BAB16379.1; -.
DR   MIM; 605666; -.
DR   InterPro; IPR004071; Cysleuk_receptor.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01533; CYSLTRECPTR.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     42       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     43    63       1 (POTENTIAL).
```

```
FT   DOMAIN       64     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     93       2 (POTENTIAL).
FT   DOMAIN       94    123       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    124    144       3 (POTENTIAL).
FT   DOMAIN      145    153       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    154    174       4 (POTENTIAL).
FT   DOMAIN      175    204       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    205    225       5 (POTENTIAL).
FT   DOMAIN      226    245       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    246    266       6 (POTENTIAL).
FT   DOMAIN      267    286       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    287    307       7 (POTENTIAL).
FT   DOMAIN      308    346       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    111    187       BY SIMILARITY.
FT   CARBOHYD     20     20       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     30     30       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    181    181       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   346 AA;  39635 MW;  EB54A4A2DDCE5EE4 CRC64;

  Query Match             17.5%;  Score 376.5;  DB 1;  Length 346;
  Best Local Similarity   30.1%;  Pred. No. 2.8e-14;
  Matches   93;  Conservative   62;  Mismatches  129;  Indels   25;  Gaps    9;

Qy         82 TITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFR 141
              :|:|    |  : |: : :| |:     :   :  | :||::|  ||  | :::: :|:
Db         11 SISVSEMEPNGTFSNNNSRNCTI----ENFKREFFPIVYLIIFFWGVLGNGLSIY-VFLQ 65

Qy        142 --TRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCS 199
                :|    ||  ||||:| ||  ||||:  |:| |:||:||:: ||   :   | ||| |
Db         66 PYKKSTSVNVFMLNLAISDLLFISTLPFRADYYLRGSNWIFGDLACRIMSYSLYVNMYSS 125

Qy        200 ILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQP 259
              |  |  :|: |:||:|||     :     | : ||::|    | |    :|
Db        126 IYFLTVLSVVRFLAMVHPFRLLHVTSIRSAWILCGIIW---ILIMASSIMLLDSGSEQNG 182

Qy        260 DITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYD------- 312
              :|:| :::      :  |   :|:|   | |:||  :  ||  ||| |   :
Db        183 SVTSCLELN--LYKIAKLQTMNYIALV-VGCLLPFFTLSICYLLIIRVLLKVEVPESGLR 239

Qy        313 --HRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYN-NTDGLYFIYLIALCLGSLN 369
                ||    :   :: |:|| :|| | : :  :|   :       | |:   :| | | : |
Db        240 VSHRKA--LTTIIITLIIFFLCFLPYHTLRTVHLTTWKVGLCKDRLHKALVITLALAAAN 297

Qy        370 SCLDPFLYF 378
              :| :| ||:
Db        298 ACFNPLLYY 306


RESULT 32
P2Y5_HUMAN
ID   P2Y5_HUMAN     STANDARD;      PRT;   344 AA.
AC   P43657; O15133;
DT   01-NOV-1995 (Rel. 32, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   P2Y purinoceptor 5 (P2Y5) (Purinergic receptor 5) (RB intron encoded
DE   G-protein coupled receptor).
GN   P2RY5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bohm S.K., Trumpp A., Khitin L.M., Kong W., Payan D.G., Bunnett N.W.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94063891; PubMed=7902321;
RA   Toguchida J., McGee T.L., Paterson J.C., Eagle J.R., Tucker S.,
RA   Yandell D.W., Dryja T.P.;
RT   "Complete genomic sequence of the human retinoblastoma susceptibility
RT   gene.";
RL   Genomics 17:535-543(1993).
RN   [3]
RP   IDENTIFICATION.
RA   Church G.M.;
RL   Unpublished observations (OCT-1994).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC       FRAMESHIFT IN POSITION 31.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L11910; -; NOT_ANNOTATED_CDS.
DR   EMBL; AF000546; AAB62190.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
KW   Palmitate.
FT   DOMAIN        1     19       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     20     46       1 (POTENTIAL).
FT   DOMAIN       47     55       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     56     79       2 (POTENTIAL).
FT   DOMAIN       80     92       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     93    112       3 (POTENTIAL).
FT   DOMAIN      113    133       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    134    154       4 (POTENTIAL).
FT   DOMAIN      155    181       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    182    209       5 (POTENTIAL).
FT   DOMAIN      210    227       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    228    253       6 (POTENTIAL).
FT   DOMAIN      254    272       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    273    292       7 (POTENTIAL).
FT   DOMAIN      293    344       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   LIPID       284    284       PALMITATE (BY SIMILARITY).
FT   DISULFID     89    168       POTENTIAL.
FT   CONFLICT     33     33       V -> I (IN REF. 2).
SQ   SEQUENCE   344 AA;  39377 MW;  7E31155AFFB499AB CRC64;

  Query Match             17.4%;  Score 376;  DB 1;  Length 344;
  Best Local Similarity   30.9%;  Pred. No. 3e-14;
  Matches   96;  Conservative   55;  Mismatches  122;  Indels   38;  Gaps   10;

Qy        101 NATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-------TTVFYTN 153
              |::  :   |    |   :: :|||:|: :| | :::      ||        || :  |
Db          5 NSSHCFYNDSFKYTLYGCMFSMVFVLGLVSNCVAIYIF------ICVLKVRNETTTYMIN 58

Qy        154 LAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLA 213
              ||::| ||  ||||:| ::    || ||::||: : ::|| ||| ||| | |||::|:||
Db         59 LAMSDLLFVFTLPFRI-FYFTTRNWPFGDLLCKISVMLFYTNMYGSILFLTCISVDRFLA 117

Qy        214 IVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCES 273
              ||:||  : |      | : |  || ||      |           ||   :  ::    |
Db        118 IVYPFKSKTLRTKRNAKIVCTGVWLTVIGGSAP-------AVFVQSTHSQGNNASEACFE 170

Qy        274 SSPFQLY--YFISLAFF----GFLIPFVLIIYCYAAIIRTLN---AYDHRWLWYVKASLL 324
              : |   :  |   :  |    || || :| : | : :::||       :   |   :
Db        171 NFPEATWKTYLSRIVIFIEIVGFFIPLILNVTCSSMVLKTLTKPVTLSRSKINKTKVLKM 230

Qy        325 I---LVIFTICFAPSNIILIIH---HANYYYNNT--DGLYFIYLIALCLGSLNSCLDPFL 376
              |   |:||  || | | ||::       : | :    :  :| | ||:   | | || :
Db        231 IFVHLIIFCFCFVPYNINLILYSLVRTQTFVNCSVVAAVRTMYPITLCIAVSNCCFDPIV 290
```

```
Qy        377 YFLMSKTRNHS 387
              |:  | |  :|
Db        291 YYFTSDTIQNS 301


RESULT 33
CLT1_PIG
ID   CLT1_PIG       STANDARD;      PRT;   340 AA.
AC   Q95N02;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 1 (CysLTR1).
GN   CYSLTR1 OR CYSLT1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takasaki J., Kamohara M., Saito T., Matsumoto M., Matsumoto S.,
RA   Ohishi T., Soga T., Matsushime H., Furuichi K.;
RT   "Characterization of cloned rat and porcine cysteinyl leukotriene
RT   receptors.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes mediating
CC       constriction of the microvascular smooth muscle during an
CC       inflammatory response. This response is mediated via a G-protein
CC       that activates a phosphatidylinositol-calcium second messenger
CC       system (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the .EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB052686; BAB60826.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     31       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     32     52       1 (POTENTIAL).
FT   DOMAIN       53     60       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     61     81       2 (POTENTIAL).
FT   DOMAIN       82    109       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    110    130       3 (POTENTIAL).
FT   DOMAIN      131    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    145    165       4 (POTENTIAL).
FT   DOMAIN      166    196       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    197    217       5 (POTENTIAL).
FT   DOMAIN      218    233       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    234    254       6 (POTENTIAL).
FT   DOMAIN      255    279       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    280    300       7 (POTENTIAL).
FT   DOMAIN      301    340       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     99    176       BY SIMILARITY.
FT   CARBOHYD      6      6       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     18     18       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    172    172       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    265    265       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   340 AA;  38986 MW;  54F9372A121CE413 CRC64;

  Query Match             17.4%;  Score 375.5;  DB 1;  Length 340;
  Best Local Similarity   31.5%;  Pred. No. 3.2e-14;
  Matches  106;  Conservative   56;  Mismatches  138;  Indels   37;  Gaps   12;

Qy         99 VKNATMGYLTSS--------LSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTT 148
              |:| |:   :|:           ::   :| :: |||    |    |::|   :  :|
Db          4 VRNLTVSCASSNTCNDTIDDFRNQVYSTLYSMITVVGFFGNGFVLYVLIKTYHEKS-AYQ 62

Qy        149 VFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISI 208
              |:  |||:|| |   ||| :: |:::   |:||: ||| :|  | |:|||   :  :|
Db         63 VYMINLAVADLLCVCTLPLRVVYYVHKGIWLFGDFLCRLSTYALYVNLYCSIFFMTAMSF 122

Qy        209 NRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVH 268
               | :||| |      |  |  | : |  :|  | |   ||  |    |  : : | | :
Db        123 FRCIAIVFPVQNINLITHKKAKIVCIAIWIFVILTSSPF--LMSTSYKDEKNNTKCFEPP 180

Qy        269 NTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTL--NAYDHRWLWYVKASLLIL 326
                ::     : :::|| | ||:||||:|| ||  || ||  |:         ||  :|:
Db        181 QXNQAKYHVLVLHYVSL-FVGFIIPFVIIIVCYTMIILTLLKNSMKKNISSRKKAIGMII 239

Qy        327 VI---FTICFAPSNIILIIHHANYYYNNT---DGLYFI---YLIALCLGSLNSCLDPFLY 377
              |:   | | | | :|   | | :: :|:|   | :  :     | |'| : | | || ||
Db        240 VVTAAFLISFMPYHIQRTI-HLHFLHNDTKHCDSVLRMQKSVXITLSLAASNCCFDPLLY 298

Qy        378 FL--------MSKTRNHSTA---YLTKXXNDLREQGQ 403
              |         :|  | || :   |: |    | |: |
Db        299 FFSGGNFREGLSTFRKHSLSTMTYVPKKKTSLPEKAQ 335


RESULT 34
CLT1_HUMAN
ID   CLT1_HUMAN     STANDARD;      PRT;   337 AA.
AC   Q9Y271;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 1 (CysLTR1) (Cysteinyl leukotriene D4
DE   receptor) (LTD4 receptor) (HG55) (HMTMF81).
GN   CYSLTR1 OR CYSLT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Tonsil;
RX   MEDLINE=99318129; PubMed=10391245;
RA   Lynch K.R., O'Neill G.P., Liu Q., Im D.-S., Sawyer N., Metters K.M.,
RA   Coulombe N., Abramovitz M., Figueroa D.J., Zeng Z., Connolly B.M.,
RA   Bai C., Austin C.P., Chateauneuf A., Stocco R., Greig G.M.,
RA   Kargman S., Hooks S.B., Hosfield E., Williams D.L. Jr.,
RA   Ford-Hutchinson A.W., Caskey C.T., Evans J.F.;
RT   "Characterization of the human cysteinyl leukotriene CysLT1
RT   receptor.";
RL   Nature 399:789-793(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peripheral blood monocytes, Leukocyte, and Spleen;
RX   MEDLINE=99393629; PubMed=10462554;
RA   Sarau H.M., Ames R.S., Chambers J., Ellis C., Elshourbagy N.,
RA   Foley J.J., Schmidt D.B., Muccitelli R.M., Jenkins O., Murdock P.R.,
RA   Herrity N.C., Halsey W., Sathe G., Muir A.I., Nuthulaganti P.,
RA   Dytko G.M., Buckley P.T., Wilson S., Bergsma D.J., Hay D.W.P.;
RT   "Identification, molecular cloning, expression, and characterization
RT   of a cysteinyl leukotriene receptor.";
RL   Mol. Pharmacol. 56:657-663(1999).
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes mediating
CC       bronchoconstriction of individuals with and without asthma.
CC       Stimulation by LTD4 results in the contraction and proliferation
CC       of smooth muscle, edema, eosinophil migration and damage to the
CC       mucus layer in the lung. This response is mediated via a G-protein
CC       that activates a phosphatidylinositol-calcium second messenger
CC       system. The rank order of affinities for the leukotrienes is LTD4
CC       >> LTE4 = LTC4 >> LTB4.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in
CC       spleen and peripheral blood leukocytes. Lower expression in
```

```
CC       several tissues, such as lung (mostly in smooth muscle bundles
CC       and alveolar macrophages), placenta, small intestine, pancreas,
CC       colon and heart.
CC   -!- MISCELLANEOUS: Selective antagonists, such as montelukast
CC       (Singulair), zafirlukast (Accolate) and pranlukast (Onon), are
CC       used in the treatment of the asthma crisis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF119711; AAD42285.1; -.
DR   EMBL; AF133266; AAD42778.1; -.
DR   HSSP; P34996; 1DDD.
DR   MIM; 300201; -.
DR   InterPro; IPR004071; Cysleuk_receptor.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01533; CYSLTRECPTR.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     28       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    29     49       1 (POTENTIAL).
FT   DOMAIN      50     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    58     78       2 (POTENTIAL).
FT   DOMAIN      79    106       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   107    127       3 (POTENTIAL).
FT   DOMAIN     128    141       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   142    162       4 (POTENTIAL).
FT   DOMAIN     163    193       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   194    214       5 (POTENTIAL).
FT   DOMAIN     215    230       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   231    251       6 (POTENTIAL).
FT   DOMAIN     252    276       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   277    297       7 (POTENTIAL).
FT   DOMAIN     298    337       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    96    173       BY SIMILARITY.
FT   CARBOHYD     6      6       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   180    180       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   262    262       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   337 AA;  38541 MW;  B9B53940F895F245 CRC64;

  Query Match             17.3%;  Score 372.5;  DB 1;  Length 337;
  Best Local Similarity   30.2%;  Pred. No. 4.6e-14;
  Matches  103;  Conservative   59;  Mismatches  140;  Indels   39;  Gaps   12;

Qy         92 ESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTV 149
              :   :| | :||          ::   :| :: |||  |    |::|   :  :|    |
Db          2 DETGNLTVSSATCHDTIDDFRNQVYSTLYSMISVVGFFGNGFVLYVLIKTYHKKS-AFQV 60

Qy        150 FYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISIN 209
              :  |||:|| |   ||| :: |:::   |:||: ||| :|   | |:|||  :  :|
Db         61 YMINLAVADLLCVCTLPLRVVYYVHKGIWLFGDFLCRLSTYALYVNLYCSIFFMTAMSFF 120

Qy        210 RYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHN 269
              | :||| |     |   |   |  :|  | |   || : | :    : : | | :
Db        121 RCIAIVFPVQNINLVTQKKARFVCVGIWIFVILTSSPFLMAKPQ--KDEKNNTKCFEPPQ 178

Qy        270 TCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTL-------NAYDHRWLWYVKAS 322
              :: :   : :::|| | ||:||||:|| ||  || ||        |   |:     ||
Db        179 DNQTKNHVLVLHYVSL-FVGFIIPFVIIIVCYTMIILTLLKKSMKKNLSSHK-----KAI 232

Qy        323 LLILVI---FTICFAPSNIILIIHHANYYYNNT---DGLYFI---YLIALCLGSLNSCLD 373
              :|:|:   | : | | :|   | | :: :| |   | :  :    :| | | : | | |
Db        233 GMIMVVTAAFLVSFMPYHIQRTI-HLHFLHNETKPCDSVLRMQKSVVITLSLAASNCCFD 291

Qy        374 PFLYFL--------MSKTRNH---STAYLTKXXNDLREQGQ 403
                 | |||         :|  | |   |  |: :    | |:|:
Db        292 PLLYFFSGGNFRKRLSTFRKHSLSSVTYVPRKKASLPEKGE 332


RESULT 35
CLT1_MOUSE
ID   CLT1_MOUSE     STANDARD;      PRT;   352 AA.
AC   Q99JA4; Q9JK47; Q9JJ71;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 1 (CysLTR1) (Cysteinyl leukotriene D4
DE   receptor) (LTD4 receptor).
GN   CYSLTR1 OR CYSLT1 OR CYSLT1R.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   STRAIN=129/Sv;
RX   PubMed=11226226;
RA   Maekawa A., Kanaoka Y., Lam B.K., Austen K.F.;
RT   "Identification in mice of two isoforms of the cysteinyl leukotriene 1
RT   receptor that result from alternative splicing.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:2256-2261(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=C57BL/6; TISSUE=Trachea;
RX   MEDLINE=21562332; PubMed=11705452;
RA   Martin V., Sawyer N., Stocco R., Unett D., Lerner M.R., Abramovitz M.,
RA   Funk C.D.;
RT   "Molecular cloning and functional characterization of murine
RT   cysteinyl-leukotriene 1 (CysLT1) receptors.";
RL   Biochem. Pharmacol. 62:1193-1200(2001).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Blood;
RX   MEDLINE=21240338; PubMed=11342226;
RA   Mollerup J., Jorgensen S.T., Hougaard C., Hoffmann E.K.;
RT   "Identification of a murine cysteinyl leukotriene receptor by
RT   expression in Xenopus laevis oocytes.";
RL   Biochim. Biophys. Acta 1517:455-459(2001).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=T-cell;
RA   Ogasawara H., Izumi T., Shimizu T.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes mediating
CC       constriction of the microvascular smooth muscle during an
CC       inflammatory response. This response is mediated via a G-protein
CC       that activates a phosphatidylinositol-calcium second messenger
CC       system. The rank order of affinities for the leukotrienes is LTD4
CC       >> LTE4 = LTC4 >> LTB4.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1/long form (shown here) and
CC       2/short form; are produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: Widely expressed, with higher expression in
CC       the lung and skin, intermediate levels in the heart, kidney and
CC       stomach and lower levels in several other tissues. Isoform 1 is
CC       the most abundant form in all tested tissues.
CC   -!- MISCELLANEOUS: MK-571, a selective antagonist, was shown to
CC       inhibit eosinophilia, bronchial hyperreactivity and microvascular
CC       leakage. Zafirlukast (Accolate) and pranlukast (Onon) were also
CC       shown to be selective antagonists.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF329272; AAK16715.1; -.
DR   EMBL; AF329272; AAK16716.1; -.
DR   EMBL; AF205830; AAK15433.1; -.
DR   EMBL; AF263370; AAF73047.1; -.
DR   EMBL; AB044087; BAA96809.1; -.
DR   HSSP; P34996; 1DDD.
DR   MGD; MGI:1926218; Cysltr1.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Alternative splicing.
FT   DOMAIN        1     43       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     44     64       1 (POTENTIAL).
FT   DOMAIN       65     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     93       2 (POTENTIAL).
FT   DOMAIN       94    121       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    122    142       3 (POTENTIAL).
FT   DOMAIN      143    156       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    157    177       4 (POTENTIAL).
FT   DOMAIN      178    208       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    209    229       5 (POTENTIAL).
FT   DOMAIN      230    245       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    246    266       6 (POTENTIAL).
FT   DOMAIN      267    291       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    292    312       7 (POTENTIAL).
FT   DOMAIN      313    352       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    111    188       BY SIMILARITY.
FT   CARBOHYD     15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     19     19       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    184    184       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     13       MISSING (IN ISOFORM 2).
FT   CONFLICT    176    176       Y -> D (IN REF. 3).
SQ   SEQUENCE   352 AA;  40715 MW;  5BDC94B3F1CD0CAB CRC64;

  Query Match             17.2%;  Score 371.5;  DB 1;  Length 352;
  Best Local Similarity   30.5%;  Pred. No. 5.3e-14;
  Matches  107;  Conservative   49;  Mismatches  120;  Indels   75;  Gaps   12;

Qy        108 TSSLSTKLI----------------PAIYLLVFVVGVPANAVTLWMLF--FRTRSICTTV 149
                  | :|:| ||                  :| :: |||   |:  |::|   :  :|    |
Db         17 TENLTTSLINNTCHDTIDEFRNQVYSTMYSVISVVGFFGNSFVLYVLIKTYHEKS-AFQV 75

Qy        150 FYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISIN 209
                :  |||||| |   ||| :: |:::   |:||: ||| ||   | |:|||  :  :|
Db         76 YMINLAIADLLCVCTLPLRVVYYVHKGKWLFGDFLCRLTTYALYVNLYCSIFFMTAMSFF 135

Qy        210 RYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHN 269
                | :||| |      |      |   | :|  | |   || : |   |  : : | | :
Db        136 RCVAIVFPVQNINLVTQKKARFVCIGIWIFVILTSSPFLMYKS--YQDEKNNTKCFEPPQ 193

Qy        270 TCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTL-------NAYDHRWLWYVKAS 322
                ::     : :::|| ||||:|||| || ||  || ||       |    |     ||
Db        194 NNQAKKYVLILHYVSL-FFGFIIPFVTIIVCYTMIILTLLKNTMKKNMPSRR-----KAI 247

Qy        323 LLILVI---FTICFAPSNIILIIHHANYYYNNTDGLYFIY----------------LIAL 363
                :|:|:   | : | | :|   ||            |: ::                :| |
Db        248 GMIIVVTAAFLVSFMPYHIQRTIH-----------LHLLHSETRPCDSVLRMQKSVVITL 296

Qy        364 CLGSLNSCLDPPLYFL--------MSKTRNHSTA---YLTKXXNDLREQGQ 403
                | : | | || |||         :|  | || :   |: |    | |:|:
Db        297 SLAASNCCFDPLLYFFSGGNFRRRLSTFRKHSLSSMTYVPKKKASLPEKGE 347


RESULT 36
P2Y4_MOUSE
ID   P2Y4_MOUSE     STANDARD;      PRT;   361 AA.
AC   Q9JJS7;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   P2Y purinoceptor 4 (P2Y4).
GN   P2RY4 OR P2Y4R.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=21185993; PubMed=11290369;
RA   Suarez-Huerta N., Pouillon V., Boeynaems J.-M., Robaye B.;
RT   "Molecular cloning and characterization of the mouse P2Y4 nucleotide
RT   receptor.";
RL   Eur. J. Pharmacol. 416:197-202(2001).
CC   -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
CC       activate a phosphatidylinositol-calcium second messenger system.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in the liver, intestine, stomach,
CC       bladder and lung.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ277752; CAB91043.1; -.
DR   HSSP; P34996; 1DDD.
DR   MGD; MGI:1926594; P2ry4.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     30       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     31     58       1 (POTENTIAL).
FT   DOMAIN       59     68       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     69     91       2 (POTENTIAL).
FT   DOMAIN       92    108       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    109    127       3 (POTENTIAL).
FT   DOMAIN      128    149       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    150    170       4 (POTENTIAL).
FT   DOMAIN      171    192       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    193    218       5 (POTENTIAL).
FT   DOMAIN      219    242       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    243    265       6 (POTENTIAL).
FT   DOMAIN      266    283       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    284    305       7 (POTENTIAL).
FT   DOMAIN      306    361       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    104    181       BY SIMILARITY.
FT   CARBOHYD    175    175       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   361 AA;  41034 MW;  3E8EA84B65BC0A20 CRC64;

  Query Match             17.2%;  Score 370;  DB 1;  Length 361;
  Best Local Similarity   31.4%;  Pred. No. 6.6e-14;
  Matches   98;  Conservative   49;  Mismatches  129;  Indels   36;  Gaps   11;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFYT-NLAIADFLFCVTLPFKIAYHL 173
                  |:|  | :|||:|:  || |||:  || |    |  |  :||::| |: ::||  : |:
Db         33 LLPLSYAVVFVLGLALNAPTLWLFLFRLRPWDATATYMFHLALSDTLYVLSLPTLVYYYA 92

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
                 -|:| ||   |:    :|| |:|||:| | |||::||: | ||       :  :| : |
Db         93 ARNHWPFGTGFCKFVRFLFYWNLYCSVLFLTCISVHRYMGICHPLRAIRWGRPRFAGLLC 152
```

```
Qy        234 GLVWATVFLYMLP--FFILKQEYYLVQPDITT--CHDVHNTCESSSPFQLYY--FISLAF 287
Db        153 LGVWLVVAGCLVPNLFFV------TTNANGTTILCHD--TTLPEEFDHYVYFSSTIMVLL 204

Qy        288 FGFLIPFVLIIYCYAAIIRTL------NAYDHRWLWYVKASLLILVIFTICFAPSNIILI 341
Db        205 FGF--PFLITLVCYGLMARRLYRPLPGAGQSSSRLRSLRTIAVVLTVFAVCFVPFHITRT 262

Qy        342 IHHANYYYNN----TDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNHSTAYLTKXXN 396
Db        263 IYYLARLLNAECRVLNIVNVVYKVTRPLASANSCLDPVLYLFTGDKYRNQ--------LQ 314

Qy        397 DLREQGQPSQRT 408
Db        315 QLCRGSTPKRRT 326
```

RESULT 37
P2Y2_HUMAN
ID   P2Y2_HUMAN   STANDARD;   PRT;   377 AA.
AC   P41231; Q96EM8;
DT   01-FEB-1995 (Rel. 31, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)
DE   (Purinergic receptor).
GN   P2RY2 OR P2RU1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Airway epithelium;
RX   MEDLINE=94211846; PubMed=8159738;
RA   Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H.,
RA   Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;
RT   "Cloning and expression of a human P2U nucleotide receptor, a target
RT   for cystic fibrosis pharmacotherapy.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:3275-3279(1994).
RN   [2]
RP   REVISIONS.
RX   MEDLINE=95108098; PubMed=7809171;
RA   Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H.,
RA   Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;
RT   "Cloning and expression of a human P2U nucleotide receptor, a target
RT   for cystic fibrosis pharmacotherapy.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:13067-13067(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Leukocyte;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
CC   activate a phosphatidylinositol-calcium second messenger system.
CC   The affinity range is UTP = ATP > ATP-gamma-S >> 2-methylthio-ATP
CC   = ADP.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: SPLEEN, TESTIS, KIDNEY, LIVER, LUNG, HEART AND
CC   BRAIN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; U07225; AAC04923.1; -.
DR   EMBL; BC012104; AAH12104.1; -.
DR   EMBL; BC028135; AAH28135.1; -.
DR   PIR; A54946; A54946.
DR   HSSP; P34996; 1DDD.
DR   Genew; HGNC:8541; P2RY2.
DR   MIM; 600041; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN   1   32   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   33   59   1 (POTENTIAL).
FT   DOMAIN   60   70   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   71   93   2 (POTENTIAL).
FT   DOMAIN   94   110   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   111   129   3 (POTENTIAL).
FT   DOMAIN   130   152   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   153   172   4 (POTENTIAL).
FT   DOMAIN   173   194   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   195   220   5 (POTENTIAL).
FT   DOMAIN   221   246   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   247   269   6 (POTENTIAL).
FT   DOMAIN   270   287   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   288   309   7 (POTENTIAL).
FT   DOMAIN   310   377   CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD   9   9   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   13   13   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID   106   183   BY SIMILARITY.
FT   CONFLICT   312   312   R -> S (IN REF. 3; AAH12104).
FT   CONFLICT   350   350   E -> G (IN REF. 1).
FT   CONFLICT   359   359   S -> F (IN REF. 1).
SQ   SEQUENCE   377 AA; 42289 MW; EE557A857A269AC6 CRC64;

  Query Match   17.0%; Score 366; DB 1; Length 377;
  Best Local Similarity   29.2%; Pred. No. 1.1e-13;
  Matches   91; Conservative   52; Mismatches 135; Indels   34; Gaps   9;

```
Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRS-ICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
Db         35 LLPVSYGVVCVLGLCLNAVALYIFLCRLKTWNASTTYMFHLAVSDALYAASLPLLVYYYA 94

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
Db         95 RGDHWPFSTVLCKLVRFLFYTNLYCSILFLTCISVHRCLGVLRPLRSLRWGRARYARRVA 154

Qy        234 GLVWATVFLYMLP--FFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFL 291
Db        155 GAVWVLVLACQAPVLYFVTTS----ARGGRVTCHDT-SAPELFSRFVAYSSVMLGLL-FA 208

Qy        292 IPFVLIIYCYAAIIR--------TLNAYDHRWLWYVKASLLILVIFTICFAPSNIILIIH 343
Db        209 VPFAVILVCYVLMARRLLKPAYGTSGGLPRAKRKSVRTIAVVLAVFALCFLPFHVTRTLY 268

Qy        344 HANYYYNNTD-------GLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXN 396
Db        269 ---YSFRSLDLSCHTLNAINMAYKVTRPLASANSCLDPVLYFLAGQR-------LVRFAR 318

Qy        397 DLREQGQPSQRT 408
Db        319 DAKPPTGPSPAT 330
```

RESULT 38
CLT1_RAT
ID   CLT1_RAT   STANDARD;   PRT;   339 AA.
AC   Q924T8;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 1 (CysLTR1).
GN   CYSLTR1 OR CYSLT1.
OS   Rattus norvegicus (Rat).

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takasaki J., Kamohara M., Saito T., Matsumoto M., Matsumoto S.,
RA   Ohishi T., Soga T., Matsushime H., Furuichi K.;
RT   "Characterization of cloned rat and porcine cysteinyl leukotriene
RT   receptors.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes mediating
CC       constriction of the microvascular smooth muscle during an
CC       inflammatory response. This response is mediated via a G-protein
CC       that activates a phosphatidylinositol-calcium second messenger
CC       system (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB052685; BAB60825.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     30       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     31     51       1 (POTENTIAL).
FT   DOMAIN       52     59       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     60     80       2 (POTENTIAL).
FT   DOMAIN       81    108       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    109    129       3 (POTENTIAL).
FT   DOMAIN      130    143       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    144    164       4 (POTENTIAL).
FT   DOMAIN      165    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    196    216       5 (POTENTIAL).
FT   DOMAIN      217    232       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    233    253       6 (POTENTIAL).
FT   DOMAIN      254    278       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    279    299       7 (POTENTIAL).
FT   DOMAIN      300    339       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     98    175       BY SIMILARITY.
FT   CARBOHYD      6      6       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    171    171       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   339 AA;  39143 MW;  281B41DF050DF8EA CRC64;

  Query Match             16.9%;  Score 365.5;  DB 1;  Length 339;
  Best Local Similarity   31.5%;  Pred. No. 1.1e-13;
  Matches  102;  Conservative   46;  Mismatches  117;  Indels   59;  Gaps   11;

Qy        119 IYLLVFVVGVPANAVTLWMLF--FRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGN 176
              :| :: |||   |:  |::|   :  :|    |:  |||||| |   ||| :: |:::
Db         31 MYSMISVVGFFGNSFVLYVLIKTYHEKS-AFQVYMINLAIADLLCVCTLPLRVVYYVHKG 89

Qy        177 NWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLV 236
              | ||: ||| ||   | |:|||  :  :|  | :||| |     |     |   | :
Db         90 KWFFGDFLCRLTTYALYVNLYCSIFFMTAMSFFRCVAIVFPVQNINLVTQKKARFVCVGI 149

Qy        237 WATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVL 296
              |- | |   || : |  |  : : | | :    ::      : :::||  |||:||||
Db        150 WIFVILTSSPFLLSKS--YQDEKNNTKCFEPPQDKQTKKYVLVLHYVSL-IFGFIIPFVT 206

Qy        297 IIYCYAAIIRTL-------NAYDHRWLWYVKASLLILVI---FTICFAPSNIILIIHHAN 346
              || ||  || ||       |    |     ||  :|:|:   | : | | :|   ||
Db        207 IIVCYTMIILTLLKNTMKKNLPSRR-----KAIGMIIVVTAAFLVSFMPYHIQRAIH--- 258

Qy        347 YYYNNTDGLYFIY---------------LIALCLGSLNSCLDPFLYFL--------MSK 382
                          |:|::                :| | | : | | || |||         :|
Db        259 --------LHFLHSETRSCDSVLRMQKSVVITLSLAASNCCFDPLLYFFSGGNFRRRLST 310

Qy        383 TRNHSTA---YLTKXXNDLREQGQ 403
                  | || :   |: |    | |:|:
Db        311 FRKHSLSSMTYIPKKKASLPEKGE 334


RESULT 39
P2Y2_RAT
ID   P2Y2_RAT       STANDARD;      PRT;   374 AA.
AC   P41232;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)
DE   (Purinergic receptor).
GN   P2RY2 OR P2RU1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95110548; PubMed=7811468;
RA   Rice W.R., Burton F.M., Fiedeldey D.T.;
RT   "Cloning and expression of the alveolar type II cell P2u-purinergic
RT   receptor.";
RL   Am. J. Respir. Cell Mol. Biol. 12:27-32(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Pituitary;
RA   Chen Z.P., Krull N., Xu S., Levy A., Lightman S.L.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar;
RX   MEDLINE=98099857; PubMed=9437211;
RA   Seye C.I., Gadeau A.P., Daret D., Dupuch F., Alzieu P., Capron L.,
RA   Desgranges C.;
RT   "Overexpression of P2Y2 purinoceptor in intimal lesions of the rat
RT   aorta.";
RL   Arterioscler. Thromb. Vasc. Biol. 17:3602-3610(1997).
CC   -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
CC       activate a phosphatidylinositol-calcium second messenger system.
CC       The affinity range is UTP = ATP > ATP-gamma-S >> 2-methylthio-ATP
CC       = ADP.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U09402; AAA61565.1; -.
DR   EMBL; L46865; AAB02099.1; -.
DR   EMBL; U56839; AAC00048.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     32       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     33     59       1 (POTENTIAL).
FT   DOMAIN       60     70       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     71     93       2 (POTENTIAL).
FT   DOMAIN       94    110       EXTRACELLULAR (POTENTIAL).
```

```
FT   TRANSMEM    111    129       3 (POTENTIAL).
FT   DOMAIN      130    152       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    153    172       4 (POTENTIAL).
FT   DOMAIN      173    194       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    195    220       5 (POTENTIAL).
FT   DOMAIN      221    245       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    246    268       6 (POTENTIAL).
FT   DOMAIN      269    286       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    287    308       7 (POTENTIAL).
FT   DOMAIN      309    374       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      9      9       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     13     13       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    106    183       BY SIMILARITY.
FT   CONFLICT    132    132       C -> S (IN REF. 1).
FT   CONFLICT    143    143       S -> R (IN REF. 1).
FT   CONFLICT    166    166       A -> T (IN REF. 1).
FT   CONFLICT    188    188       A -> D (IN REF. 1).
SQ   SEQUENCE   374 AA;  42137 MW;  2DA6853AA20A1AF3 CRC64;

  Query Match             16.9%;  Score 365.5;  DB 1;  Length 374;
  Best Local Similarity   30.2%;  Pred. No. 1.2e-13;
  Matches   93;  Conservative   52;  Mismatches  130;  Indels   33;  Gaps   10;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRS-ICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              |:|  | :| |:|:  | | |::   | ::   :| :  :||::| |:  :||  : |:
Db         35 LLPVSYGVVCVLGLCLNVVALYIFLCRLKTWNASTTYMFHLAVSDSLYAASLPLLVYYYA 94

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
              |::| |  |||:    :|| |:||||| | |||::| | :: |         ||
Db         95 QGDHWPFSTVLCKLVRFLFYTNLYCSILFLTCISVHRCLGVLRPLHSLSWGHARYARRVA 154

Qy        234 GLVWATVFLYMLP--FFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFL 291
              :||  |     |  :|:       |:    ||||  :  |  | |  |  : |    |
Db        155 AVVWVLVLACQAPVLYFVTTS----VRGTRITCHDT-SARELFSHFVAYSSVMLGLL-FA 208

Qy        292 IPFVLIIYCYAAIIRTL--NAYDHRWL-----WYVKASLLILVIFTICFAPSNIILIIHH 344
              :|| :|: ||  : | |   ||    |       |:   |:| :| :|| | ::   ::
Db        209 VPFSIILVCYVLMARRLLKPAYGTTGLPRAKRKSVRTIALVLAVFALCFLPFHVTRTLY- 267

Qy        345 ANYYYNNTD-------GLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXND 397
              | : : |       :   | |   | | |||||| ||||  :        | :   |
Db        268 --YSFRSLDLSCHTLNAINMAYKITRPLASANSCLDPVLYFLAGQR-------LVRFARD 318

Qy        398 LREQGQPS 405
              :   :|:
Db        319 AKPATEPT 326


RESULT 40
GP34_HUMAN
ID   GP34_HUMAN     STANDARD;      PRT;   381 AA.
AC   Q9UPC5; O95853;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable G protein-coupled receptor GPR34.
GN   GPR34.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=99326137; PubMed=10395919;
RA   Schoneberg T., Schulz A., Grosse R., Schade R., Henklein P.,
RA   Schultz G., Gudermann T.;
RT   "A novel subgroup of class I G-protein-coupled receptors.";
RL   Biochim. Biophys. Acta 1446:57-70(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99156852; PubMed=10036181;
RA   Marchese A., Sawzdargo M., Nguyen T., Cheng R., Heng H.H., Nowak T.,
RA   Im D.-S., Lynch K.R., George S.R., O'Dowd B.F.;
RT   "Discovery of three novel orphan G-protein-coupled receptors.";
RL   Genomics 56:12-21(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20434921; PubMed=10982042;
RA   Jacobi F.K., Broghammer M., Pesch K., Zrenner E., Berger W.,
RA   Meindl A., Pusch C.M.;
RT   "Physical mapping and exclusion of GPR34 as the causative gene for
RT   congenital stationary night blindness type 1.";
RL   Hum. Genet. 107:89-91(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ORPHAN RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- TISSUE SPECIFICITY: BROADLY EXPRESSED.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039686; AAD50531.1; -.
DR   EMBL; AF118670; AAD17248.1; -.
DR   EMBL; AK027780; BAB55362.1; -.
DR   EMBL; BC020678; AAH20678.1; -.
DR   Genew; HGNC:4490; GPR34.
DR   MIM; 300241; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor; G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     61       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     62     82       1 (POTENTIAL).
FT   DOMAIN       83     88       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     89    109       2 (POTENTIAL).
FT   DOMAIN      110    128       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    129    149       3 (POTENTIAL).
FT   DOMAIN      150    171       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    172    192       4 (POTENTIAL).
FT   DOMAIN      193    216       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    217    237       5 (POTENTIAL).
FT   DOMAIN      238    269       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    270    290       6 (POTENTIAL).
FT   DOMAIN      291    310       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    311    331       7 (POTENTIAL).
FT   DOMAIN      332    381       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    127    204       BY SIMILARITY.
FT   CARBOHYD     28     28       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     36     36       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    200    200       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    181    181       L -> V (IN REF. 1).
```

```
SQ   SEQUENCE   381 AA;  43860 MW;  491FC016E5624379 CRC64;

  Query Match             16.9%;  Score 365.5;  DB 1;  Length 381;
  Best Local Similarity   30.4%;  Pred. No. 1.2e-13;
  Matches  102;  Conservative   53;  Mismatches  131;  Indels   49;  Gaps   12;

Qy        101 NATMGYLTSSLSTKLIPAI----YLLVFVVGVPANAVTLWM-LFFRTRSICTTVFYTNLA 155
              :||    |  :  ||:  :    | ::|:||:  | : |:: |    :     ::  |:|
Db         38 SATPNVTTCPMDEKLLSTVLTTSYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYLLNVA 97

Qy        156 IADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIV 215
              ||| |    |||:| ||:| | |  | :||:    :|| ||| ||:||  ||::||: |
Db         98 IADLLLIFCLPFRIMYHINQNKWTLGVILCKVVGTLFYMNMYISIILLGFISLDRYIKIN 157

Qy        216 HPFTYRGLPKHTYALVTCGLVWATV---FLYMLPFFILKQEYYLVQPDITTC---HDVHN 269
                    |       ::  | :||      || |:   : |  :     : | |    | ||
Db        158 RSIQQRKAITTKQSIYVCCIVWMLALGGFLTMIILTLKKGGH-----NSTMCFHYRDKHN 212

Qy        270 TCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHR---------WLWYVK 320
                  :    :: || :  |  :  |:|||  |  | : |     |         :    :
Db        213 AKGEA----IFNFILVVMFWLI--FLLIILSYIKIGKNLLRISKRRSKFPNSGKYATTAR 266

Qy        321 ASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYL------IALCLGSLNSCLDP 374
              | ::|:|||||| | :    |: ::    |    |:  :       | | | | ||||||
Db        267 NSFIVLIIFTICFVPYHAFRFIYISSQL--NVSSCYWKEIVHKTNEIMLVLSSFNSCLDP 324

Qy        375 FLYFLMSKTRNHSTAYLTKXXNDL---REQGQPSQ 406
               :|||||         : |    |   | ||:||:
Db        325 VMYFLMSSN-------IRKIMCQLLFRRFQGEPSR 352


Search completed: June 24, 2003, 12:03:47
Job time : 17.5178 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:59:21 ; Search time 26.0319 Seconds
(without alignments)
1506.722 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIKM.....AYLTKXXNDLREQGQPSQRT 408

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	606.5	28.1	399	2	I48705	proteinase activat
2	599.5	27.8	420	2	I51667	thrombin receptor
3	589.5	27.3	397	2	S66518	proteinase-activat
4	523	24.2	425	2	A37912	thrombin receptor
5	505.5	23.4	432	2	A43448	thrombin receptor
6	501.5	23.2	427	2	S17148	alpha-thrombin rec
7	424	19.7	362	2	S33733	G protein-coupled
8	410.5	19.0	361	2	B45680	G protein-coupled
9	406.5	18.8	308	2	I50241	G protein-coupled
10	399	18.5	373	2	JC4737	G protein-coupled
11	398	18.5	373	2	JC4162	P2Y receptor - bov
12	384.5	17.8	370	2	JC5549	heptahelical P2Y5-
13	382	17.7	365	2	S68679	G protein-coupled
14	376	17.4	344	2	T09508	intron 17 purinerg
15	363.5	16.9	355	2	JQ1231	interleukin-8 rece
16	361	16.7	373	2	A47556	ATP receptor P2u -
17	356.5	16.5	362	2	JN0694	angiotensin II rec
18	356	16.5	375	2	A54946	P-2U nucleotide re
19	354	16.4	350	2	A39445	interleukin-8 rece
20	353	16.4	356	2	S42096	interleukin-8 rece
21	352	16.3	328	2	I55450	G protein-coupled
22	351.5	16.3	360	2	A53611	interleukin-8 rece
23	351	16.3	365	2	S68208	G protein-coupled
24	347	16.1	361	2	JC5653	G protein-coupled
25	346	16.0	341	2	S63666	platelet activatin
26	345.5	16.0	359	2	A48921	interleukin-8 rece
27	345.5	16.0	362	2	B57641	G protein-coupled
28	345	16.0	359	2	S15403	angiotensin II rec
29	344	15.9	372	2	S26667	G protein-coupled
30	343	15.9	359	2	I51372	angiotensin II rec
31	342.5	15.9	342	2	A40191	platelet-activatin
32	342.5	15.9	358	2	A53752	interleukin-8 rece
33	342	15.9	374	2	S32785	G protein-coupled
34	342	15.9	380	2	A48227	kappa opioid recep
35	341.5	15.8	342	2	S13638	platelet-activatin
36	341.5	15.8	369	2	JC5068	G protein-coupled
37	341	15.8	359	2	A48857	angiotensin II rec
38	341	15.8	359	2	JC2134	angiotensin II rec
39	341	15.8	380	2	JC2434	kappa opioid recep
40	340	15.8	341	2	S43252	platelet-activatin
41	340	15.8	359	2	JH0621	angiotensin II rec
42	340	15.8	380	2	S36143	kappa opioid recep
43	339.5	15.7	354	2	I53033	G protein-coupled
44	339.5	15.7	362	2	A57641	G protein-coupled
45	339	15.7	353	2	JC2492	G protein-coupled

ALIGNMENTS

RESULT 1
I48705
proteinase activated receptor 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C;Accession: I48705
R;Nystedt, S.; Larsson, A.K.; Aberg, H.; Sundelin, J.
J. Biol. Chem. 270, 5950-5955, 1995
A;Title: The mouse proteinase-activated receptor-2 cDNA and gene. Molecular cloning and
A;Reference number: I48705; MUID:95197620; PMID:7890726
A;Accession: I48705
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-399 <RES>
A;Cross-references: EMBL:Z48043; NID:g663020; PIDN:CAA88097.1; PID:g663021
C;Superfamily: ATP receptor P2u

Query Match 28.1%; Score 606.5; DB 2; Length 399;
Best Local Similarity 35.4%; Pred. No. 2.9e-43;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

```
Qy         22 ALIFAAAGLLLLLPTFCQSGMEN--DTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEGWT 79
              :| :   |: ||  :   |  ||     | :|    |      ||            |
Db          5 SLAWLLGGITLLAASVSCSRTENLAPGRNNSKGRSLIGRLETQPP------------IT 51

Qy         80 GATITVK--IKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWM 137
              |  : |:      | |||            ||  |:|  :| :|::|||:|:|:| : ||:
Db         52 GKGVPVEPGFSIDEFSAS----------ILTGKLTTVFLPVVYIIVFVIGLPSNGMALWI 101

Qy        138 LFFRTRSICTTVFY-TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNM 196
               |||:     | |  |||:|| |   :  | ||:|||:|||||:|| ||:     |||||
Db        102 FLFRTKKKHPAVIYMANLALADLLSVIWFPLKISYHLHGNNWVYGEALCKVLIGFFYGNM 161

Qy        197 YCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKQE 253
              ||||| : |:|: ||  ||:|    | |:   |  :   |:   :|  :||  :| :::||
Db        162 YCSILFMTCLSVQRYWVIVNPM---GHPRKK-ANIAVGVSLAIWLLIFLVTIPLYVMKQT 217

Qy        254 YYLVQPDITTCHDVHNTCESSSPFQLY------YFISLAFFGFLIPFVLIIYCYAAIIRT 307
               |:   :|||||||         | ::       ||:|||   || | :|    |. :|:|
Db        218 IYIPALNITTCHDV-------LPEEVLVGDMFNYFLSLAIGVFLFPALLTASAYVLMIKT 270

Qy        308 LNA------YDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLI 361
              | :       : :    ::  : :| :: |||||||::|::|:           :| :||:
Db        271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNLLLVVHYFLIKTQRQSHVYALYLV 330

Qy        362 ALCLGSLNSCLDPFLYFLMSKT-RNHS 387
              |||| :||||:|||:|: :||  |:|:
Db        331 ALCLSTLNSCIDPFVYYFVSKDFRDHA 357
```

RESULT 2
I51667

```
thrombin receptor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I51667
R;Gerszten, R.E.; Chen, J.; Ishii, M.; Ishii, K.; Nanevicz, T.; Turck, C.W.; Vu, T.H.; C
Nature 368, 648-651, 1994
A;Title: Thrombin receptor's specificity for agonist peptide is determined by its extrac
A;Reference number: I51667; MUID:94195429; PMID:8145852
A;Accession: I51667
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-420 <GER>
A;Cross-references: EMBL:U09632; NID:g495197; PIDN:AAA18498.1; PID:g495198

  Query Match           27.8%;  Score 599.5;  DB 2;  Length 420;
  Best Local Similarity  34.6%;  Pred. No. 1.2e-42;
  Matches  128;  Conservative   68;  Mismatches  149;  Indels   25;  Gaps    8;

Qy         30 LLLLLPT-------FCQSGMENDTNNLAKPTLPIKTFR--GAPPNSFEEFPFSAL-EGWT 79
Db          8 LLLLLLTLLGAMGSLCLANSDTQAKGAHSNNMTIKTFRIFDDSESEFEEIPWDELDESGE 67

Qy         80 GATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLF 139
Db         68 GSGDQAPVSRSARKPIRRNITKEAEQYLSSQWLTKFVPSLYTVVFIVGLPLNLLAIIIFL 127

Qy        140 FRTRSICTTVFY-TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYC 198
Db        128 FKMKVRKPAVVYMLNLAIADVFFVSVLPFKIAYHLSGNDWLFGPGMCRIVTAIFYCNMYC 187

Qy        199 SILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQ 258
Db        188 SVLLIASISVDRFLAVVYPMHSLSWRTMSRAYMACSFIWLISIASTIPLLVTEQTQKIPR 247

Qy        259 PDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYD------ 312
Db        248 LDITTCHDVLD-LKDLKDFYIYYFSSFCLLFFFVPFIITTICYIGIIRSLSSSSIENSCK 306

Qy        313 -HRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSC 371
Db        307 KTRALF---LAVVVLCVFIICFGPTNVLFLTH---YLQEANEFLYFAYILSACVGSVSCC 360

Qy        372 LDPFLYFLMS 381
Db        361 LDPLIYYYAS 370


RESULT 3
S66518
proteinase-activated receptor 2 precursor - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Nov-2000
C;Accession: S66518; S64709; G02131
R;Nystedt, S.; Emilsson, K.; Larsson, A.K.; Stroembeck, B.; Sundelin, J.
Eur. J. Biochem. 232, 84-89, 1995
A;Title: Molecular cloning and functional expression of the gene encoding the human prot
A;Reference number: S66518; MUID:96048032; PMID:7556175
A;Accession: S66518
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <NYS>
A;Cross-references: EMBL:Z49993; NID:g1008084; PIDN:CAA90290.1; PID:g1008085
R;Boehm, S.K.; Kong, W.; Broemme, D.; Smeekens, S.P.; Anderson, D.C.; Connolly, A.; Kahn
Biochem. J. 314, 1009-1016, 1996
A;Title: Molecular cloning, expression and potential functions of the human proteinase-a
A;Reference number: S64709; MUID:96177879; PMID:8615752
A;Accession: S64709
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-137,'A',139-397 <BOE>
A;Cross-references: EMBL:U34038; NID:g1041728; PIDN:AAB47871.1; PID:g1041729
A;Note: the authors translated the codon GTC for residue 68 as Ile and AAC for residue 2
R;Kahn, M.L.; Coughlin, S.R.
submitted to the EMBL Data Library, September 1995
A;Reference number: H00822
A;Accession: G02131
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 29-397 <KAH>
A;Cross-references: EMBL:U36753; NID:g1208539; PIDN:AAA90957.1; PID:g1208540
C;Genetics:
A;Map position: 5q13
A;Introns: 28/1
C;Superfamily: ATP receptor P2u
F;1-36/Domain: activation peptide #status predicted <APT>
F;1-25/Domain: signal sequence #status predicted <SIG>
F;37-397/Product: proteinase-activated receptor 2 #status predicted <MAT>

  Query Match           27.3%;  Score 589.5;  DB 2;  Length 397;
  Best Local Similarity  35.0%;  Pred. No. 7.8e-42;
  Matches  137;  Conservative   65;  Mismatches  138;  Indels   51;  Gaps   13;

Qy         28 AGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEGWTGATITVKI 87
Db         12 AAILLAASLSC-SGTIQGTNRSSKGRSLIGKVDGT----------SHVTG-KGVTVETVF 59

Qy         88 KCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICT 147
Db         60 SVDEFSAS----------VLTGKLTTVFLPIVYTIVFVVGLPSNGMALWVFLFRTKKKHP 109

Qy        148 TVFY-TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACI 206
Db        110 AVIYMANLALADLLSVIWFPLKIAYHIHGNNWIYGEALCNVLIGFFYGNMYCSILFMTCL 169

Qy        207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHD 266
Db        170 SVQRYWVIVNPMGH-SRKKANIAIGISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHD 228

Qy        267 VHNTCESSSPFQLY------YFISLAFFGFLIPFVLIIYCYAAIIRTLNA------YDHR 314
Db        229 V-------LPEQLLVGDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENSEKK 281

Qy        315 WLWYVKASLLILVIFTICFAPSNIILIIHHANYYYNNTDG---LYFIYLIALCLGSLNSC 371
Db        282 RKRAIKLIVTVLAMYLICFTPSNLLLVVH---YFLIKSQGQSHVYALYIVALCLSTLNSC 338

Qy        372 LDPFLYFLMSKT-RNHS-TAYLTKXXNDLRE 400
Db        339 IDPFVYYFVSHDFRDHAKNALLCRSVRTVKQ 369


RESULT 4
A37912
thrombin receptor precursor - human
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 05-Nov-1999
C;Accession: A37912
R;Vu, T.K.H.; Hung, D.T.; Wheaton, V.I.; Coughlin, S.R.
Cell 64, 1057-1068, 1991
A;Title: Molecular cloning of a functional thrombin receptor reveals a novel proteolyti
A;Reference number: A37912; MUID:91168254; PMID:1672265
A;Accession: A37912
A;Molecule type: mRNA
A;Residues: 1-425 <VUA>
A;Cross-references: GB:M62424; NID:g339676; PIDN:AAA36743.1; PID:g339677
C;Genetics:
A;Gene: GDB:F2R
A;Cross-references: GDB:127737; OMIM:187930
A;Map position: 5q13-5q13
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-425/Product: thrombin receptor #status predicted <MAT>

  Query Match           24.2%;  Score 523;  DB 2;  Length 425;
  Best Local Similarity  32.9%;  Pred. No. 3.1e-36;
  Matches  127;  Conservative   76;  Mismatches  153;  Indels   30;  Gaps   11;
```

```
Qy          14 GTQVIKMKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPN-SFEEFPF 72
Db           2 GPRRLLLVAACFSLCGPLLSARTRARRPESKATN----ATLDPRSFLLRNPNDKYEPFWE 57

Qy          73 SALEGWTGAT----ITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGV 128
Db          58 DEEKNESGLTEYRLVSINKSSPLQKQLPAFISEDASGYLTSSWLTLFVPSVYTGVFVVSL 117

Qy         129 PAN--AVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCR 186
Db         118 PLNIMAIVVFILKMKVKK-PAVVYMLHLATADVLFVSVLPFKISYYFSGSDWQFGSELCR 176

Qy         187 ATTVIFYGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
Db         177 FVTAAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASF---TCLAIWALAIAG 233

Qy         244 MLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAA 303
Db         234 VVPLVLKEQTIQVPGLNITTCHDVLNETLLEG-YYAYYFSAFSAVFFFVPLIISTVCYVS 292

Qy         304 IIRTLN-------AYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANY-YYNNTDGL 355
Db         293 IIRCLSSSAVANRSKKSRALF---LSAAVFCIFIICFGPTNVLLIAHYSFLSHTSTTEAA 349

Qy         356 YFIYLIALCLGSLNSCLDPFLYFLMS 381
Db         350 YFAYLLCVCVSSISSCIDPLIYYYAS 375
```

RESULT 5
A43448
thrombin receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: A43448
R;Zhong, C.; Hayzer, D.J.; Corson, M.A.; Runge, M.S.
J. Biol. Chem. 267, 16975-16979, 1992
A;Title: Molecular cloning of the rat vascular smooth muscle thrombin receptor. Evidence
A;Reference number: A43448; MUID:92381002; PMID:1324917
A;Accession: A43448
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-432 <ZHO>
A;Cross-references: GB:M81642; NID:g207465; PIDN:AAA42274.1; PID:g207466
A;Experimental source: RASM aortic smooth muscle cells
A;Note: sequence extracted from NCBI backbone (NCBIN:111973, NCBIP:111974)
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match 23.4%; Score 505.5; DB 2; Length 432;
  Best Local Similarity 31.2%; Pred. No. 9.1e-35;
  Matches 134; Conservative 71; Mismatches 163; Indels 61; Gaps 15;

```
Qy          21 KALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKT-------FRGAPPNSFEEFPFS 73
Db           4 RRLLLVAVGLSLCGPLLSSRVPMRQPESERMYATPYATPNPRSFFLRNPSEDTFEQFPLG 63

Qy          74 ---------ALEGWTGATITVKIKCP--------EESASHLHVKNATMGYLTSSLSTKLI 116
Db          64 DEEEKNESIPLEG--RAVYLNKSRFPPMPPPPFISEDAS---------GYLTSPWLTLFI 112

Qy         117 PAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFY-TNLAIADFLFCVTLPFKIAYHLNG 175
Db         113 PSVYTFVFIVSLPLNILAIAVFVFRMKVKKPAVVYMLHLAMADVLFVSVLPFKISYYFSG 172

Qy         176 NNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVT 232
Db         173 TDWQFGSGMCRFATAACYCNMYASIMLMTVISIDRFLAVVYPIQSLSWRTLGRANF---T 229

Qy         233 CGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLI 292
Db         230 CVVIWVMAIMGVVPLLLKEQTTQVPGLNITTCHDVLNETLLHG-FYSYYFSAFSAIFFLV 288

Qy         293 PFVLIIYCYAAIIRTLN-------AYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHA 345
Db         289 PLIISTVCYTSIIRCLSSSAVANRSKKSRALF---LSAAVFCIFIVCFGPTNVLLIVHYL 345

Qy         346 NYYYN-NTDGLYFIYLIALCLGSLNSCLDPFL-YFLMSKTRNHSTAYL----TKXXNDLR 399
Db         346 LLSDSPGTETAYFAYLLCVCVTSVASCIDPLIYYYASSECQKHLYSILCCRESSDSNSCN 405

Qy         400 EQGQ--PSQ 406
Db         406 STGQLMPSK 414
```

RESULT 6
S17148
alpha-thrombin receptor - Chinese hamster
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C;Accession: S17148
R;Rasmussen, U.B.; Vouret-Craviari, V.; Jallat, S.; Schlesinger, Y.; Pages, G.; Paviran
FEBS Lett. 288, 123-128, 1991
A;Title: cDNA cloning and expression of a hamster alpha-thrombin receptor coupled to Ca
A;Reference number: S17148; MUID:91348247; PMID:1652467
A;Accession: S17148
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-427 <RAS>
A;Cross-references: EMBL:X61958; NID:g940495; PIDN:CAA43957.1; PID:g49538
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match 23.2%; Score 501.5; DB 2; Length 427;
  Best Local Similarity 32.1%; Pred. No. 2e-34;
  Matches 126; Conservative 73; Mismatches 140; Indels 53; Gaps 13;

```
Qy          23 LIFAAAGLLLLLPTFC--------QSGMENDTNNLAKPTLPIKTF-RGAPPNSFEEFPFS 73
Db           6 LLLVAAGLSLCGPLLSSRVPVRQPESEMTDATVN------PRSFFLRNPGENTFELIPLG 59

Qy          74 ALEGWTGATITVKIKCPEESASHLH----------VKNATMGYLTSSLSTKLIPAIYLL 122
Db          60 DEEEKNESTL------PEGRAIYLNKSHSPAPLAPFISEDASGYLTSPWLRLFIPSVYTF 113

Qy         123 VFVVGVPAN--AVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVF 180
Db         114 VFVVSLPLNILAIAVFVLKMKVKK-PAVVYMLHLAMADVLFVSVLPLKISYYFSGSDWQF 172

Qy         181 GEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVW 237
Db         173 GSGMCRFATAAFYCNMYASIMLMTVISIDRFLAVVYPIQSLSWRTLGRANF---TCLVIW 229

Qy         238 ATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLI 297
Db         230 VMAIMGVVPLLLKEQTTRVPGLNITTCHDVLNETLLQG-FYSYYFSAFSAVFFLVPLIIS 288

Qy         298 IYCYAAIIRTLNAYD-------HRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYYN 350
Db         289 TICYMSIIRCLSSSSVANRSKKSRALF---LSAAVFCVFIVCFGPTNVLLIMHYLLLSDS 345

Qy         351 -NTDGLYFIYLIALCLGSLNSCLDPFLYFLMS 381
Db         346 PATEKAYFAYLLCVCVSSVSCCIDPLIYYYAS 377
```

RESULT 7
S33733
G protein-coupled receptor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
C;Accession: S33733
R;Webb, T.E.; Simon, J.; Krishek, B.J.; Bateson, A.N.; Smart, T.G.; King, B.F.; Burnstc
FEBS Lett. 324, 219-225, 1993
A;Title: Cloning and functional expression of a brain G-protein-coupled ATP receptor.
A;Reference number: S33733; MUID:93285340; PMID:8508924

```
A;Accession: S33733
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-362 <WEB>
A;Cross-references: EMBL:X73268; NID:g395084; PIDN:CAA51716.1; PID:g395085
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             19.7%;  Score 424;  DB 2;  Length 362;
  Best Local Similarity   28.9%;  Pred. No. 5.1e-28;
  Matches  101;  Conservative   59;  Mismatches  143;  Indels   46;  Gaps    9;

Qy          77 GWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLW 136
               ||     | |     :      :                  :| :|:|||: |   |:| :|
Db          21 GWAAGNATTKCSLTKTGFQFYY----------------LPTVYILVFITGFLGNSVAIW 63

Qy         137 MLFFRTRSIC-TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGN 195
               |  |  |     :|:  |||:|||| :|||  | |: |  :|:||:|:|:     ||: |
Db          64 MFVFHMRPWSGISVYMFNLALADFLYVLTLPALIFYYFNKTDWIFGDVMCKLQRFIFHVN 123

Qy         196 MYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY 255
               :| ||| | |||::||  :|||    |  |    |:    ||||  |   : |       :|
Db         124 LYGSILFLTCISVHRYTGVVHPLKSLGRLKKKNAVYVSSLVWALVVAVIAPIL-----FY 178

Qy         256 ----LVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAY 311
                   : :    ||:|.  | :         | :    | | |||::|: ||  |:: |
Db         179 SGTGVRRNKTITCYD--TTADEYLRSYFVYSMCTTVFMFCIPFIVILGCYGLIVKALIYK 236

Qy         312 D-------HRWLWYVKASLLILVIFTICFAPSNIILIIH-HANYYYNN------TDGLYF 357
               |       : :: |   :::| :| : : | :::  ::  |   :         | :|
Db         237 DLDNSPLRRKSIYLV---IIVLTVFAVSYLPFHVMKTLNLRARLDFQTPQMCAFNDKVYA 293

Qy         358 IYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
                | :   | |||||:|| ||||   |     :  |:  :   |    |:
Db         294 TYQVTRGLASLNSCVDPILYFLAGDTFRRRLSRATRKSSRRSEPNVQSK 342


RESULT 8
B45680
G protein-coupled peptide receptor EBI 2 - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: B45680
R;Birkenbach, M.; Josefsen, K.; Yalamanchili, R.; Lenoir, G.; Kieff, E.
J. Virol. 67, 2209-2220, 1993
A;Title: Epstein-Barr virus-induced genes: first lymphocyte-specific G protein-coupled p
A;Reference number: A45680; MUID:93188173; PMID:8383238
A;Accession: B45680
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-361 <BIR>
A;Cross-references: GB:L08177; NID:g292056; PIDN:AAA35924.1; PID:g292057
A;Experimental source: B-lymphocytes
A;Note: sequence extracted from NCBI backbone (NCBIN:127096, NCBIP:127097)
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             19.0%;  Score 410.5;  DB 2;  Length 361;
  Best Local Similarity   30.2%;  Pred. No. 6.8e-27;
  Matches   99;  Conservative   58;  Mismatches  114;  Indels   57;  Gaps   10;

Qy         106 YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSI-CTTVFYTNLAIADFLFCVT 164
               |    | :  ::|  | |||::|:  | : | ::     : |  ||:: ||| |:| ||
Db          24 YAHHSTARIVMPLHYSLVFIIGLVGNLLALVVIVQNRKKINSTTLYSTNLVISDILFTTA 83

Qy         165 LPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLP 224
               || :|||:  | :|  |: ||| | ::|| | |  :  : |:||:|::|:|||  |  :
Db          84 LPTRIAYYAMGFDWRIGDALCRITALVFYINTYAGVNFMTCLSIDRFIAVVHPLRYNKIK 143

Qy         225 KHTYALVTCGLVWATVFLYMLPFFI---LKQEYYLVQPDITTCHDVHNTCESSS-PFQLY 280
               :  :|   |  ||  ||   ||  |    |||      :  || :  |  |: | |: |
Db         144 RIEHAKGVCIFVWILVFAQTLPLLINPMSKQE-----AERITCMEYPNFEETKSLPWIL- 197

Qy         281 YFISLAFFGFLIPFVLIIYCYAAII-------------------RTLNAYDHRWLWYVK 320
                  :   | |:::| ::|: ||: |                       : ||
Db         198 --LGACFIGYVLPLIIILICYSQICCKLFRTAKQNPLTEKSGVNKKALNTI--------- 246

Qy         321 ASLLILVIFTICFAPSNIILIIH------HANYYYNNTDGLYFIYL-IALCLGSLNSCLD 373
                  :||:|:| :|| | :: :| |       :|:   :    : | |   :|| : | |:|
Db         247 --ILIIVVFVLCFTPYHVAIIQHMIKKLRFSNFLECSQRHSFQISLHFTVCLMNFNCCMD 304

Qy         374 PFLYFLMSKTRNHSTAYLTKXXNDLREQ 401
               ||:||   |       |  |    |: |
Db         305 PFIYFFACK------GYKRKVMRMLKRQ 326


RESULT 9
I50241
G protein-coupled receptor 6H1 - chicken
N;Alternate names: purinoceptor 6H1
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Jun-2000
C;Accession: I50241; JC4618
R;Kaplan, M.H.; Smith, D.I.; Sundick, R.S.
J. Immunol. 151, 628-636, 1993
A;Title: Identification of a G protein coupled receptor induced in activated T cells.
A;Reference number: I50241; MUID:93329058; PMID:8393036
A;Accession: I50241
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-308 <KAP>
A;Cross-references: GB:L06109; NID:g304383; PIDN:AAB06587.1; PID:g304384
R;Webb, T.E.; Kaplan, M.G.; Barnard, E.A.
Biochem. Biophys. Res. Commun. 219, 105-110, 1996
A;Title: Identification of 6H1 as a P2Y purinoceptor: P2Y5.
A;Reference number: JC4618; MUID:96190677; PMID:8619790
A;Accession: JC4618
A;Molecule type: mRNA
A;Residues: 1-308 <WEB>
A;Cross-references: GB:L06109; NID:g304383; PIDN:AAB06587.1; PID:g304384
A;Experimental source: T-cells
C;Comment: This receptor plays a role in T-cell activation.
C;Genetics:
A;Gene: p2Y5
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein
F;15-40/Domain: transmembrane #status predicted <TM1>
F;51-74/Domain: transmembrane #status predicted <TM2>
F;89-109/Domain: transmembrane #status predicted <TM3>
F;133-153/Domain: transmembrane #status predicted <TM4>
F;177-201/Domain: transmembrane #status predicted <TM5>
F;227-248/Domain: transmembrane #status predicted <TM6>
F;269-292/Domain: transmembrane #status predicted <TM7>

  Query Match             18.8%;  Score 406.5;  DB 2;  Length 308;
  Best Local Similarity   33.5%;  Pred. No. 1.3e-26;
  Matches  108;  Conservative   52;  Mismatches  127;  Indels   35;  Gaps   11;

Qy         107 LTSSLSTK------LIPAIYLLVFVVGVPANAVTLWMLFF--RTRSICTTVFYTNLAIAD 158
               ::|: ||:       |   :: :|||:|: || | :::  |  : |:  || :  ||||:|
Db           2 VSSNCSTEDSFKYTLYGCVFSMVFVLGLIANCVAIYIFTFTLKVRNE-TTTYMLNLAISD 60

Qy         159 FLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF 218
               ||  ||||:| |:    || ||:|||: :  :|| ||| ||| | |||::|:|||||||
Db          61 LLFVFTLPFRI-YYFVVRNWPFGDVLCKISVTLFYTNMYGSILFLTCISVDRFLAIVHPF 119

Qy         219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESS-SPF 277
               : |      | : |  || ||     |    :        :  || :  |  ||:   :
Db         120 RSKTLRTKRNARIVCVAVWITVLAGSTPASFFQSTNRQNNTEQRTCFE--NFPESTWKTY 177

Qy         278 QLYYFISLAFFGFLIPFVLIIYCYAAIIRTLN---AYDHRWLWYVKASLLI---LVIFTI 331
                  | :   || || :| : |   ::|||          |   |   :|   ||||
Db         178 LSRIVIFIEIVGFFIPLILNVTCSTMVLRTLNKPLTLSRNKLSKKKVLKMIFVHLVIFCF 237

Qy         332 CFAPSNIILIIH---HANYYYNNT--DGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNH 386
               || | || ||::       : | :    :  :| : ||:   | | || :|:  | |
```

```
Db        238 CFVPYNITLILYSLMRTQTWINCSVVTAVRTMYPVTLCIAVSNCCFDPIVYYFTSDTN-- 295

Qy        387 STAYLTKXXNDLREQGQPSQRT 408
                  ::| :: |  | |
Db        296 ---------SELDKKQQVHQNT 308


RESULT 10
JC4737
G protein-coupled receptor P2Y1 - human
N;Alternate names: P2Y1 purinergic receptor; P2Y1 purinoceptor
C;Species: Homo sapiens (man)
C;Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 17-Nov-2000
C;Accession: JC4737; JC4615; S54253
R;Janssens, R.; Communi, D.; Pirotton, S.; Samson, M.; Parmentier, M.; Boeynaems, J.M.
Biochem. Biophys. Res. Commun. 221, 588-593, 1996
A;Title: Cloning and tissue distribution of the human P2Y1 receptor.
A;Reference number: JC4737; MUID:96205320; PMID:8630005
A;Accession: JC4737
A;Molecule type: DNA
A;Residues: 1-373 <JAN>
A;Cross-references: GB:S81950; NID:g1839438; PIDN:AAB47091.1; PID:g1839439
R;Ayyanathan, K.; Webbs, T.E.; Sandhu, A.K.; Athwal, R.S.; Barnard, E.A.; Kunapuli, S.P.
Biochem. Biophys. Res. Commun. 218, 783-788, 1996
A;Title: Cloning and chromosomal localization of the human P2Y1 purinoceptor.
A;Reference number: JC4615; MUID:96158962; PMID:8579591
A;Accession: JC4615
A;Molecule type: mRNA
A;Residues: 1-373 <AYY>
A;Cross-references: GB:U42029; NID:g1147730; PIDN:AAA97872.1; PID:g1147731
A;Experimental source: erythro leukemia cells
R;Leon, C.; Vial, C.; Cazenave, J.; Gachet, C.
submitted to the EMBL Data Library, May 1995
A;Description: Cloning of a human putative P2Y receptor.
A;Reference number: S54253
A;Accession: S54253
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-137,139-373 <LEO>
A;Cross-references: EMBL:Z49205; NID:g798835; PIDN:CAA89066.1; PID:g798836
C;Comment: This receptor belongs to a family of G protein-coupled receptors. It responds
C;Genetics:
A;Gene: p2Y1; GDB:P2RY1
A;Cross-references: GDB:677125; OMIM:601167
A;Map position: 3pter-3qter
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F;52-77/Domain: transmembrane #status predicted <TM1>
F;88-111/Domain: transmembrane #status predicted <TM2>
F;124-152/Domain: transmembrane #status predicted <TM3>
F;171-191/Domain: transmembrane #status predicted <TM4>
F;214-237/Domain: transmembrane #status predicted <TM5>
F;261-282/Domain: transmembrane #status predicted <TM6>
F;305-328/Domain: transmembrane #status predicted <TM7>
F;11,27,113,197/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;258,336/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predict
F;330,339/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predict
F;343/Binding site: phosphate (Ser) (covalent) (by protein kinase C and calmodulin-depen

  Query Match           18.5%;  Score 399;  DB 2;  Length 373;
  Best Local Similarity  29.7%;  Pred. No. 6.5e-26;
  Matches  113;  Conservative   66;  Mismatches  143;  Indels   58;  Gaps   14;

Qy         52 PTLPIKT---FRGAPPNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLT 108
              | :|  |   |   | :|           |  :|:          ||   | ::    ||
Db          7 PAVPNGTDAAFLAGPGSS----------WGNSTV----------ASTAAVSSSFKCALT 45

Qy        109 -SSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYTNLAIADFLFCVTLP 166
               :      :||:|:|||::|   |:| :||  |  :     :|:  |||:|||| : :|||
Db         46 KTGFQFYYLPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVLTLP 105

Qy        167 FKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKH 226
                | |: |  :|:||: :|:    ||: |:| ||| | ||| :||  :|:|    |  |

Db        106 ALIFYYFNKTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKK 165

Qy        227 TYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDIT-TCHDVHNTCESSSPFQLYYFI-- 283
               |:    |||  | : : |  ||       |: : | ||:|      :|  :   |||
Db        166 KNAICISVLVWLIVVVAISP--ILFYSGTGVRKNKTITCYDT-----TSDEYLRSYFIYS 218

Qy        284 ---SLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKASLLILVIFTICF 333
                 ::|   | :| |||: ||  |:| |   |          : :: |   :::| :| : :
Db        219 MCTTVAM--FCVPLVLILGCYGLIVRALIYKDLDNSPLRRKSIYLV---IIVLTVFAVSY 273

Qy        334 APSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNH 386
               | :::  ::  |   :          | :|  | :   | |||||:|| ||||   |
Db        274 IPFHVMKTMNLRARLDFQTPAMCAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRR 333

Qy        387 STAYLTKXXNDLREQGQPSQ 406
               :  |:  :   |    |:
Db        334 RLSRATRKASRRSEANLQSK 353


RESULT 11
JC4162
P2Y receptor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 12-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999
C;Accession: JC4162
R;Henderson, D.J.; Elliot, D.G.; Smith, G.M.; Webb, T.E.; Dainty, I.A.
Biochem. Biophys. Res. Commun. 212, 648-656, 1995
A;Title: Cloning and characterisation of a bovine P2Y receptor.
A;Reference number: JC4162; MUID:95352058; PMID:7626079
A;Accession: JC4162
A;Molecule type: mRNA
A;Residues: 1-373 <HEN>
A;Cross-references: EMBL:X87628; NID:g1032484; PIDN:CAA60958.1; PID:g1032485
A;Experimental source: aortic endothelial cell
C;Genetics:
A;Gene: bovp2y
C;Superfamily: ATP receptor P2u
C;Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein
F;52-77/Domain: transmembrane #status predicted <TM1>
F;88-111/Domain: transmembrane #status predicted <TM2>
F;124-150/Domain: transmembrane #status predicted <TM3>
F;171-191/Domain: transmembrane #status predicted <TM4>
F;214-237/Domain: transmembrane #status predicted <TM5>
F;261-282/Domain: transmembrane #status predicted <TM6>
F;305-328/Domain: transmembrane #status predicted <TM7>
F;11,27,113,197/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;258/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

  Query Match           18.5%;  Score 398;  DB 2;  Length 373;
  Best Local Similarity  28.5%;  Pred. No. 7.8e-26;
  Matches  111;  Conservative   66;  Mismatches  136;  Indels   76;  Gaps   15;

Qy         45 DTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEGWTGATI--TVKIKCPEESASHLHVKNA 102
              ||  || |         |:|             |  :|:  |  :  | : |
Db         14 DTAFLADP--------GSP-------------WGNSTVTSTAAVASPFKCA-------- 43

Qy        103 TMGYLT-SSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYTNLAIADFL 160
                 || :     :||:|:|||::|   |:| :||  |  :     :|:  |||:||||
Db         44 ----LTKTGFQFYYLPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFL 99

Qy        161 FCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTY 220
               : :|||  | |: |  :|:||: :|:    ||: |:| ||| | ||| :||  :|:|
Db        100 YVLTLPALIFYYFNKTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKS 159

Qy        221 RGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY----LVQPDITTCHDVHNTCESSSP 276
               |  |   |:    |||  | : : |      :|    : :    ||:|       :|
Db        160 LGRLKKKNAVYISVLVWLIVVVGISPIL-----FYSGTGIRKNKTITCYDT-----TSDE 209

Qy        277 FQLYYFI-----SLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKASLL 324
               :   |||     ::|   | :| |||: ||  |:| |   |         : :: |   ::
Db        210 YLRSYFIYSMCTTVAM--FCVPLVLILGCYGLIVRALIYKDLDNSPLRRKSIYLV---II 264

Qy        325 ILVIFTICFAPSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDPFLY 377
```

```
Db        265 VLTVFAVSYIPFHVMKTMNLRARLDFQTPEMCAFNDRVYATYQVTRGLASLNSCVDPILY 324

Qy        378 FLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
Db        325 FLAGDTFRRRLSRATRKASRRSEANLQSK 353
```

RESULT 12
JC5549
heptahelical P2Y5-like receptor - human
C;Species: Homo sapiens (man)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 24-Sep-1999
C;Accession: JC5549
R;Janssens, R.; Boeynaems, J.M.; Godart, M.; Communi, D.
Biochem. Biophys. Res. Commun. 236, 106-112, 1997
A;Title: Cloning of a human heptahelical receptor closely related to the P2Y5 receptor.
A;Reference number: JC5549; MUID:97366605; PMID:9223435
A;Accession: JC5549
A;Molecule type: DNA
A;Residues: 1-370 <JAN>
A;Cross-references: DDBJ:AF005419; NID:g2240034; PIDN:AAB66322.1; PID:g2240035
C;Superfamily: ATP receptor P2u

  Query Match 17.8%; Score 384.5; DB 2; Length 370;
  Best Local Similarity 30.4%; Pred. No. 1.1e-24;
  Matches 95; Conservative 60; Mismatches 121; Indels 37; Gaps 10;

```
Qy         91 EESASHLHVK--NATMG---YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSI 145
Db         12 QDSNSSLRPRLGNATANNTCIVDDSFKYNLNGAVYSVVFILGLITNSVSLFVFCFRMKMR 71

Qy        146 C-TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLA 204
Db         72 SETAIFITNLAVSDLLFVCTLPFKIFYNFN-RHWPFGDTLCKISGTAFLTNIYGSMLFLT 130

Qy        205 CISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTC 264
Db        131 CISVDRFLAIVYPFRSRTIRTRRNSAIVCAGVWILVLSGGISASLF---------STTNV 181

Qy        265 HDVHNTCESSSPFQLY--YFISLAFF----GFLIPFVLIIYCYAAIIR------TLNAYD 312
Db        182 NNATTTCFEGLSKRVWKTYLSKITIFIEVVGFIIPLILNVSCSSVVLRTLRKPATLSQIG 241

Qy        313 HRWLWYVKASLLILVIFTICFAPSNIILIIHH-------ANYYYNNTDGLYFIYLIALCL 365
Db        242 TNKKKVLKMITVHMAVFVVCFVPYNSVLFLYALVRSQAITNCFLERFAKI--MYPITLCL 299

Qy        366 GSLNSCLDPFLYF 378
Db        300 ATLNCCFDPFIYY 312
```

RESULT 13
S68679
G protein-coupled receptor - human
C;Species: Homo sapiens (man)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S68679
R;Stam, N.J.; Klomp, J.; van de Heuvel, M.; Olijve, W.
FEBS Lett. 384, 260-264, 1996
A;Title: Molecular cloning and characterization of a novel orphan receptor (P(2P)) expre
A;Reference number: S68679; MUID:96197801; PMID:8617367
A;Accession: S68679
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-365 <STA>
A;Cross-references: EMBL:X96597; NID:g1296631; PIDN:CAA65415.1; PID:g1296632
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor

  Query Match 17.7%; Score 382; DB 2; Length 365;
  Best Local Similarity 31.8%; Pred. No. 1.7e-24;
  Matches 100; Conservative 47; Mismatches 127; Indels 40; Gaps 10;

```
Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFYT-NLAIADFLFCVTLPFKIAYHL 173
Db         37 LLPVSYAVVFVLGLGLNAPTLWLFIFRLRPWDATATYMFHLALSDTLYVLSLPTLIYYYA 96

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
Db         97 AHNHWPFGTEICKFVRFLFYWNLYCSVLFLTCISVHRYLGICHPLRALRWGRPRLAGLLC 156

Qy        234 GLVWATVFLYMLP--FFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFL 291
Db        157 LAVWLVVAGCLVPNLFFVTTSN----KGTTVLCHDTTRPEE----FDHYVHFSSAVMGLL 208

Qy        292 --IPFVLIIYCYAAIIRTL------NAYDHRWLWYVKASLLILVIFTICFAPSNIILIIH 343
Db        209 FGVPCLVTLVCYGLMARRLYQPLPGSAQSSSRLRSLRTIAVVLTVFAVCFVPFHITRTIY 268

Qy        344 H------ANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXND 397
Db        269 YLARLLEADCRVLNI--VNVVYKVTRPLASANSCLDPVLYLLTG----------DKYRRQ 316

Qy        398 LRE---QGQPSQRT 408
Db        317 LRQLCGGGKPQPRT 330
```

RESULT 14
T09508
intron 17 purinergic receptor P2Y5 - human
N;Alternate names: G-protein coupled receptor
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C;Accession: T09508
R;Bohm, S.K.; Trumpp, A.; Khitin, L.M.; Kong, W.; Payan, D.G.; Bunnett, N.W.
submitted to the EMBL Data Library, April 1997
A;Description: The human purinergic receptor P2Y5 is encoded in intron 17 of the retino
A;Reference number: Z16705
A;Accession: T09508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-344 <BOH>
A;Cross-references: EMBL:AF000546; NID:g2232068; PID:g2232069
C;Genetics:
A;Map position: 13
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match 17.4%; Score 376; DB 2; Length 344;
  Best Local Similarity 30.9%; Pred. No. 5e-24;
  Matches 96; Conservative 55; Mismatches 122; Indels 38; Gaps 10;

```
Qy        101 NATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-------TTVFYTN 153
Db          5 NSSHCFYNDSFKYTLYGCMFSMVFVLGLVSNCVAIYIF------ICVLKVRNETTTYMIN 58

Qy        154 LAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLA 213
Db         59 LAMSDLLFVFTLPFRI-FYFTTRNWPFGDLLCKISVMLFYTNMYGSILFLTCISVDRFLA 117

Qy        214 IVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCES 273
Db        118 IVYPFKSKTLRTKRNAKIVCTGVWLTVIGGSAP-------AVFVQSTHSQGNNASEACFE 170

Qy        274 SSPFQLY--YFISLAFF----GFLIPFVLIIYCYAAIIRTLN---AYDHRWLWYVKASLL 324
Db        171 NFPEATWKTYLSRIVIFIEIVGFFIPLILNVTCSSMVLKTLTKPVTLSRSKINKTKVLKM 230

Qy        325 I---LVIFTICFAPSNIILIIH---HANYYYNNT--DGLYFIYLIALCLGSLNSCLDPFL 376
Db        231 IFVHLIIFCFCFVPYNINLILYSLVRTQTFVNCSVVAAVRTMYPITLCIAVSNCCFDPIV 290

Qy        377 YFLMSKTRNHS 387
```

```
                |:  | |  :|
Db        291 YYFTSDTIQNS 301


RESULT 15
JQ1231
interleukin-8 receptor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Nov-1999
C;Accession: JQ1231; A46483
R;Beckmann, M.P.; Munger, W.E.; Kozlosky, C.; VandenBos, T.; Price, V.; Lyman, S.; Gerar
Biochem. Biophys. Res. Commun. 179, 784-789, 1991
A;Title: Molecular characterization of the interleukin-8 receptor.
A;Reference number: JQ1231; MUID:91378994; PMID:1898400
A;Accession: JQ1231
A;Molecule type: DNA
A;Residues: 1-355 <BEC>
A;Cross-references: GB:M74240; NID:g165438; PIDN:AAA31375.1; PID:g165439
R;Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
J. Immunol. 148, 1261-1264, 1992
A;Title: Characterization of complementary DNA clones encoding the rabbit IL-8 receptor.
A;Reference number: A46483; MUID:92148149; PMID:1737938
A;Accession: A46483
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-355 <LEE>
A;Cross-references: GB:M82873; NID:g165440; PIDN:AAA31376.1; PID:g165441
A;Experimental source: neutrophils
A;Note: sequence extracted from NCBI backbone (NCBIN:81526, NCBIP:81530)
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             16.9%;  Score 363.5;  DB 2;  Length 355;
  Best Local Similarity   29.8%;  Pred. No. 5.8e-23;
  Matches  100;  Conservative   66;  Mismatches  101;  Indels   69;  Gaps   14;

Qy        107 LTSSLSTKLIPAIYLLVFVVGVPANA-VTLWMLFFRTRSICTTVFYTNLAIADFLFCVTL 165
              :| :|:  ::  || |||:: :  |: | | :|: |:    | |:  |||:|| || :|:
Db         38 VTQTLNKYVVVVIYALVFLLSLLGNSLVMLVILYSRSNRSVTDVYLLNLAMADLLFALTM 97

Qy        166 PFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPK 225
              |         |  |:||  ||:  :::   | |  ||||||||::|||||||        :
Db         98 PIWAVSKEKG--WIFGTPLCKVVSLVKEVNFYSGILLLACISVDRYLAIVHATRTLTQKR 155

Qy        226 HTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITT--CHD--VHNTCESSSPFQLYY 281
              |    : | :||   :  ||||: :| :    |: ::  |::   ||| :    :::
Db        156 HLVKFICLG-IWALSLILSLPFFLFRQVF---SPNNSSPVCYEDLGHNTAK----WRMVL 207

Qy        282 FISLAFFGFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSN 337
              |    |||::| :::::||   :|||         || :  : |   :::|| :|: | |
Db        208 RILPHTFGFILPLLVMLFCYGFTLRTLFQAHMGQKHRAMRVIFA---VVLIFLLCWLPYN 264

Qy        338 IILIIHHANYYYNNTDGLYFIYLI------------AL----CLGSLNSCLDPFLYF--- 378
              ::|:            | |   ::|             ||     || |:|||:| :|
Db        265 LVLL----------ADTLMRTHVIQETCQRRNDIDRALDATEILGFLHSCLNPIIYAFIG 314

Qy        379 ---------------LMSK---TRNHSTAYLTKXXN 396
                            |:||   ||:  |:| :   |
Db        315 QNFRNGFLKMLAARGLISKEFLTRHRVTSYTSSSTN 350


RESULT 16
A47556
ATP receptor P2u - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C;Accession: A47556
R;Lustig, K.D.; Shiau, A.K.; Brake, A.J.; Julius, D.
Proc. Natl. Acad. Sci. U.S.A. 90, 5113-5117, 1993
A;Title: Expression cloning of an ATP receptor from mouse neuroblastoma cells.
A;Reference number: A47556; MUID:93281707; PMID:7685114
A;Accession: A47556
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-373 <LUS>
A;Cross-references: GB:L14751; NID:g309457; PIDN:AAA39871.1; PID:g309458
C;Superfamily: ATP receptor P2u
C;Keywords: transmembrane protein

  Query Match             16.7%;  Score 361;  DB 2;  Length 373;
  Best Local Similarity   29.1%;  Pred. No. 9.8e-23;
  Matches   90;  Conservative   53;  Mismatches  132;  Indels   34;  Gaps    9;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRS-ICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              |:|  | :| |:|:  | | |::   | ::   :| :  :||::| |:  :||  : |:
Db         35 LLPVSYGVVCVLGLCLNVVALYIFLCRLKTWNASTTYMFHLAVSDSLYAASLPLLVYYYA 94

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
              |::| |  |||:    :|| |:|||| | |||::| | :: |       :  ||
Db         95 RGDHWPFSTVLCKLVRFLFYTNLYCSILFLTCISVHRCLGVLRPLHSLRWGRARYARRVA 154

Qy        234 GLVWATVFLYMLP--FFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFL 291
              . :|| |     | :|:       |:    |||| :  | | |  |  : |    |
Db        155 AVVWVLVLACQAPVLYFVTTS----VRGTRITCHDT-SARELFSHFVAYSSVMLGLL-FA 208

Qy        292 IPFVLIIYCYAAIIR--------TLNAYDHRWLWYVKASLLILVIFTICFAPSNIILIIH 343
              :|| :|: ||  : |        |            |:   |:| :| :|| | ::   ::
Db        209 VPFSVILVCYVLMARRLLKPAYGTTGGLPRAKRKSVRTIALVLAVFALCFLPFHVTRTLY 268

Qy        344 HANYYYNNTD-------GLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXN 396
                | : : |         :   | |   | | |||||| ||||  :         | :
Db        269 ---YSFRSLDLSCHTLNAINMAYKITRPLASANSCLDPVLYFLAGQR-------LVRFAR 318

Qy        397 DLREQGQPS 405
              | :   :|:
Db        319 DAKPPTEPT 327


RESULT 17
JN0694
angiotensin II receptor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C;Accession: JN0694; S47627
R;Ji, H.; Sandberg, K.; Zhang, Y.; Catt, K.J.
Biochem. Biophys. Res. Commun. 194, 756-762, 1993
A;Title: Molecular cloning, sequencing and functional expression of an amphibian angiot
A;Reference number: JN0694; MUID:93343933; PMID:7688227
A;Accession: JN0694
A;Molecule type: mRNA
A;Residues: 1-362 <JIH>
A;Cross-references: GB:L16463; NID:g387890; PIDN:AAA49647.1; PID:g387891
R;Nishimatsu, S.; Koyasu, N.; Sugaya, T.; Ohnishi, J.; Yamagishi, T.; Murakami, K.; Miy
Biochim. Biophys. Acta 1218, 401-407, 1994
A;Title: Isolation and characterization of two alternatively spliced complementary DNAs
A;Reference number: S47627; MUID:94325348; PMID:7519446
A;Accession: S47627
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-362 <NIS>
A;Cross-references: GB:S73274; NID:g625009; PIDN:AAC60749.1; PID:g625010
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; phosphoprotein; transmembrane protein
F;34-54/Domain: transmembrane #status predicted <TM1>
F;66-86/Domain: transmembrane #status predicted <TM2>
F;105-125/Domain: transmembrane #status predicted <TM3>
F;146-166/Domain: transmembrane #status predicted <TM4>
F;192-212/Domain: transmembrane #status predicted <TM5>
F;241-260/Domain: transmembrane #status predicted <TM6>
F;285-305/Domain: transmembrane #status predicted <TM7>
F;3,18,177/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;222,343/Binding site: phosphate (Thr) (covalent) #status predicted
F;347/Binding site: phosphate (Ser) (covalent) #status predicted

  Query Match             16.5%;  Score 356.5;  DB 2;  Length 362;
  Best Local Similarity   30.3%;  Pred. No. 2.3e-22;
```

```
  Matches   88;  Conservative   51;  Mismatches  122;  Indels   29;  Gaps    7;

Qy        116 IPAIYLLVFVVGVPANAVTLWMLF-FRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLN 174
              || ||  :|||||  |:: : ::: :       ::|  |||::|  | :|||   ||
Db         32 IPIIYSTIFVVGVFGNSMVVIVIYSYMKMKTVASIFLMNLALSDLCFVITLPLWAAYTAM 91

Qy        175 GNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCG 234
              :| ||  ||:  :     |:| :: || |:||:|| |||||   |       | :||
Db         92 HYHWPFGNFLCKVASTAITLNLYTTVFLLTCLSIDRYSAIVHPMKSRIWRTAMVARLTCV 151

Qy        235 LVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLA----FFGF 290
              :|   ||  :|  | :| |     : | |  |:::        :|: : ::      ||
Db        152 GIWLVAFLASMPSIIYRQIYLFHDTNQTVCAIVYDS------GHIYFMVGMSLAKNIVGF 205

Qy        291 LIPFVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNIIL------ 340
              ||||::|:  |  | :||     |   |     |  : ::::|  |: |  :
Db        206 LIPFLIILTSYTLIGKTLKEVYRAQRARNDDIFKMIVAVVLLFFFCWIPYQVFTFLDVLI 265

Qy        341 ---IIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSKT-RNH 386
                :| :  |     |:     | :|:  ||||:||||    |  |.|
Db        266 QMDVIQNCKMYDIVDTGM----PITICIAYFNSCLNPFLYGFFGKNFRKH 311


RESULT 18
A54946
P-2U nucleotide receptor - human
C;Species: Homo sapiens (man)
C;Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 17-Mar-1999
C;Accession: A54946
R;Parr, C.E.; Sullivan, D.M.; Paradiso, A.M.; Lazarowski, E.R.; Burch, L.H.; Olsen, J.C.
Proc. Natl. Acad. Sci. U.S.A. 91, 3275-3279, 1994
A;Title: Cloning and expression of a human P-2U nucleotide receptor, a target for cystic
A;Reference number: A54946; MUID:94211846; PMID:8159738
A;Accession: A54946
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-375 <PAR>
A;Cross-references: GB:U07225
A;Note: parts of this sequence were confirmed by protein sequencing
C;Genetics:
A;Gene: GDB:P2RY2; HP2U; P2U
A;Cross-references: GDB:362713; OMIM:600041
A;Map position: 11q13.5-11q14.1
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match           16.5%;  Score 356;  DB 2;  Length 375;
  Best Local Similarity  29.4%;  Pred. No. 2.6e-22;
  Matches   92;  Conservative   49;  Mismatches  134;  Indels   38;  Gaps   10;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRS-ICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              |:|  | :| |:|:  ||| |::   | ::   :| :  :||::| |:  :||  : |:
Db         35 LLPVSYGVVCVLGLCLNAVGLYIFLCRLKTWNASTTYMFHLAVSDALYAASLPLLVYYYA 94

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
              |::| |  |||:    :|| |:||||| | |||::| | :: |       :  ||
Db         95 RGDHWPFSTVLCKLVRFLFYTNLYCSILFLTCISVHRCLGVLRPLRSLRWGRARYARRVA 154

Qy        234 GLVWATVFLYMLPFFILKQEYYLVQPDI---TTCHDVHNTCESSSPFQLYYFISLAFFGF 290
              | ||  |     |        | |        ||||  :  |  | |  |  : |    |
Db        155 GAVWVLVLACQAPVL------YFVTTSARGPLTCHDT-SAPELFSRFVAYSSVMLGLL-F 206

Qy        291 LIPFVLIIYCYAAIIR--------TLNAYDHRWLWYVKASLLILVIFTICFAPSNIILII 342
              :|| :|: ||  : |         |          |:  ::| :| :|| | ::   :
Db        207 AVPFAVILVCYVLMARRLLKPAYGTSGGLPRAKRKSVRTIAVVLAVFALCFLPFHVTRTL 266

Qy        343 HHANYYYNNTD-------GLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXX 395
              :   | : : |       :   | :   | | |||||| ||||  :        | :
Db        267 Y---YSFRSLDLSCHTLNAINMAYKVTR-LASANSCLDPVLYFLAGQR-------LVRFA 315

Qy        396 NDLREQGQPSQRT 408
              | :   ||  |
Db        316 RDAKPPTGPSPAT 328
```

```
RESULT 19
A39445
interleukin-8 receptor type A - human
N;Alternate names: interleukin-8 receptor, high-affinity
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1993 #sequence_revision 12-Apr-1996 #text_change 05-Nov-1999
C;Accession: I37449; I38710; I38711; A39445
R;Mollereau, C.; Muscatelli, F.; Mattei, M.G.; Vassart, G.; Parmentier, M.
Genomics 16, 248-251, 1993
A;Title: The high-affinity interleukin 8 receptor gene (IL8RA) maps to the 2q33-q36 reg
A;Reference number: I37449; MUID:93252387; PMID:8486366
A;Accession: I37449
A;Molecule type: DNA
A;Residues: 1-350 <RES>
A;Cross-references: EMBL:X65858; NID:g312046; PIDN:CAA46688.1; PID:g312047
R;Ahuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
J. Biol. Chem. 269, 26381-26389, 1994
A;Title: Comparison of the genomic organization and promoter function for human interle
A;Reference number: I37898; MUID:95014476; PMID:7929358
A;Accession: I38710
A;Molecule type: DNA
A;Residues: 1-350 <RE2>
A;Cross-references: EMBL:U11870; NID:g511804; PIDN:AAA64378.1; PID:g511805
A;Accession: I38711
A;Molecule type: mRNA
A;Residues: 1-16 <RE3>
A;Cross-references: EMBL:U11871; NID:g511806; PIDN:AAA64379.1; PID:g733002
R;Holmes, W.E.; Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
Science 253, 1278-1280, 1991
A;Title: Structure and functional expression of a human interleukin-8 receptor.
A;Reference number: A39445; MUID:91368199; PMID:1840701
A;Accession: A39445
A;Molecule type: mRNA
A;Residues: 1-275,'T',277-350 <HOL>
A;Cross-references: GB:M68932; NID:g186369; PIDN:AAA59159.1; PID:g186370
C;Genetics:
A;Gene: GDB:IL8RA
A;Cross-references: GDB:135039; OMIM:146929
A;Map position: 2q35-2q35
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; membrane protein

  Query Match           16.4%;  Score 354;  DB 2;  Length 350;
  Best Local Similarity  31.4%;  Pred. No. 3.5e-22;
  Matches   94;  Conservative   58;  Mismatches  109;  Indels   38;  Gaps   11;

Qy        108 TSSLSTKLIPAIYLLVFVVGVPANA-VTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLP 166
              | :|:  ::   | |||:: :  |: | | :|: |     | |:  |||:|| || :|||
Db         34 TETLNKYVVIIAYALVFLLSLLGNSLVMLVILYSRVGRSVTDVYLLNLALADLLFALTLP 93

Qy        167 FKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKH 226
              |  :||  |:||  ||:  :::   | |  |||||||::|||||||        :|
Db         94 IWAASKVNG--WIFGTFLCKVVSLLKEVNFYSGILLLACISVDRYLAIVHATRTLTQKRH 151

Qy        227 TYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITT--CHDVHNTCESSSPFQLYYFIS 284
              |  |  |       ||||: :| |:   |: ::  |::|      :: :::   |
Db        152 LVKFVCLG-CWGLSMNLSLPFFLFRQAYH---PNNSSPVCYEVLG--NDTAKWRMVLRIL 205

Qy        285 LAFFGFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNIIL 340
              |||::|  ::::||   :|||         || :  : |   :::|| :|: | |::|
Db        206 PHTFGFIVPLFVMLFCYGFTLRTLFKAHMGQKHRAMRVIFA---VVLIFLLCWLPYNLVL 262

Qy        341 ---------IIHHANYYYNNTDGLYFIYLIAL----CLGSLNSCLDPFLYFLMSKTRNH 386
                      :|  :   ||          ||      || |:|||:| :|  : :   |
Db        263 LADTLMRTQVIQESCERRNNIGR-------ALDATEILGFLHSCLNPIIYAFIGQNFRH 314


RESULT 20
S42096
interleukin-8 receptor - rat
C;Species: Rattus norvegicus (Norway rat)
```

```
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999
C;Accession: S42096
R;Gobl, A.E.; Wang, S.; Zhou, Y.; Oeberg, K.
submitted to the EMBL Data Library, February 1994
A;Description: Molecular cloning of the rat IL8 receptor.
A;Reference number: S42096
A;Accession: S42096
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-356 <GOB>
A;Cross-references: EMBL:X77797
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             16.4%;  Score 353;  DB 2;  Length 356;
  Best Local Similarity   30.3%;  Pred. No. 4.4e-22;
  Matches   95;  Conservative   59;  Mismatches  116;  Indels   44;  Gaps   11;

Qy        116 IPAIYLLVFVVGVPANAVTLWMLFFRTRSIC--TTVFYTNLAIADFLFCVTLPFKIAYHL 173
              :  ||:|| :: :  |:: : :: :  || |  | |:  ||||||  | :|||  |  :
Db         50 VVVIYVLVTLLSLVGNSLVMLVILY-NRSTCSVTDVYLLNLAIADLFFALTLPVWAASKV 108

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
              ||  |:||  ||:  : :     | |:|||||||::|||||||  :     :|    | |
Db        109 NG--WIFGSFLCKVFSFLQEITFYSSVLLLACISMDRYLAIVHATSTLIQKRHLVKFV-C 165

Qy        234 GLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIP 293
              :|    :  || |||:       |    |::  |  ::| :::   |    :|||:|
Db        166 ITMWFLSLVLSLPIFILRTT-VKANPSTVVCYE--NIGNNTSKWRVVLRILPQTYGFLLP 222

Qy        294 FVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANYYY 349
              :::::||   :|||       || :  : |   ::::| :|: | ||:|
Db        223 LLIMLFCYGFTLRTLFKAHMGQKHRAMRVIFA---VVLVFLLCWLPYNIVLF-------- 271

Qy        350 NNTDGLYFIYLI------------ALCLGSLNSCLDPFLYFLMSKTRNHS-----TAYL 391
                || |    ||            :  || |:|||:| :|  : :   |        |
Db        272 --TDTLMRTKLIKETCERQNEINKASEILGFLHSCLNPIIYAFIGQKFRHGLLKIMANYG 329

Qy        392 TKXXNDLREQGQPS 405
                  | ::|:||
Db        330 LVSKEFLAKEGRPS 343


RESULT 21
I55450
G protein-coupled P2 receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
C;Accession: I55450
R;Chang, K.; Hanaoka, K.; Kumada, M.; Takuwa, Y.
J. Biol. Chem. 270, 26152-26158, 1995
A;Title: Molecular cloning and functional analysis of a novel P2 nucleotide receptor.
A;Reference number: I55450; MUID:96064682; PMID:7592819
A;Accession: I55450
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <RES>
A;Cross-references: GB:D63665; NID:g1066007; PIDN:BAA09816.1; PID:g1066008
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor

  Query Match             16.3%;  Score 352;  DB 2;  Length 328;
  Best Local Similarity   31.4%;  Pred. No. 4.9e-22;
  Matches   96;  Conservative   42;  Mismatches  126;  Indels   42;  Gaps   10;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFYT-NLAIADFLFCVTLPFKIAYHL 173
              |:| :| :| |||:| |   :  :    |::  :  || |||:|| |:  :||  |  :
Db         28 LLPPVYSVVLVVGLPLNVCVIAQICASRRTLTRSAVYTLNLALADLLYACSLPLLIYNYA 87

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKH-----TY 228
              |::| ||:: ||    :|| |:: ||| | |||  ||| | ||      | |
Db         88 RGDHWPFGDLACRLVRFLFYANLHGSILFLTCISFQRYLGICHPLA----PWHKRGGRRA 143

Qy        229 ALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPF----QLYYFIS 284
              | | ||:||  |    ||  :       :| : | |:|:      | |        | | ::
Db        144 AWVVCGVVWLVVTAQCLPTAVFAATG--IQRNRTVCYDL------SPPILSTRYLPYGMA 195

Qy        285 LAFFGFLIPFVLIIYCYAAIIRTLNAYD--------HRWLWYVKASLLILVIFTICFAPS 336
              |   |||:||  :: ||  : | |   |          |     : ::::  :| | | |
Db        196 LTVIGFLLPFTALLACYCRMARRLCRQDGPAGPVAQERRSKAARMAVVVAAVFVISFLPF 255

Qy        337 NIILIIHHANYYYNNTDGLYFIYLIALC--------LGSLNSCLDPFL-YFLMSKTRNHS 387
              :|    : |    :| |:    |               | || ||| | ||   | |
Db        256 HITKTAYLA---VRSTPGVSCPVLETFAAAYKGTRPFASANSVLDPILFYFTQQKFRRQP 312

Qy        388 TAYLTK 393
                | |
Db        313 HDLLQK 318


RESULT 22
A53611
interleukin-8 receptor type B - human
C;Species: Homo sapiens (man)
C;Date: 07-Oct-1994 #sequence_revision 12-Apr-1996 #text_change 05-Nov-1999
C;Accession: I37898; I38712; A53611; A39446
R;Ahuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
J. Biol. Chem. 269, 26381-26389, 1994
A;Title: Comparison of the genomic organization and promoter function for human interl
A;Reference number: I37898; MUID:95014476; PMID:7929358
A;Accession: I37898
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <RES>
A;Cross-references: EMBL:U11869; NID:g511801; PIDN:AAB60656.1; PID:g511803
A;Accession: I38712
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15 <RE2>
A;Cross-references: EMBL:U11872; NID:g511808; PIDN:AAA64380.1; PID:g511809; EMBL:U11873
11876; NID:g511816; PID:g511817; EMBL:U11877; NID:g511818; PID:g511819; EMBL:U11878; NI
R;Sprenger, H.; Lloyd, A.R.; Lautens, L.L.; Bonner, T.I.; Kelvin, D.J.
J. Biol. Chem. 269, 11065-11072, 1994
A;Title: Structure, genomic organization, and expression of the human interleukin-8 re
A;Reference number: A53611; MUID:94209273; PMID:7512557
A;Accession: A53611
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 6-360 <SPR>
A;Cross-references: GB:M99412; GB:L19593
R;Murphy, P.M.; Tiffany, H.L.
Science 253, 1280-1283, 1991
A;Title: Cloning of complementary DNA encoding a functional human interleukin-8 recepto
A;Reference number: A39446; MUID:91368200; PMID:1891716
A;Accession: A39446
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 6-360 <MUR>
A;Cross-references: GB:M73969
C;Comment: This receptor, unlike IL8RA, binds several peptides besides interleukin-8, i
C;Genetics:
A;Gene: GDB:IL8RB; IL8RA
A;Cross-references: GDB:127868; OMIM:146928
A;Map position: 2q35-2q35
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             16.3%;  Score 351.5;  DB 2;  Length 360;
  Best Local Similarity   27.5%;  Pred. No. 5.9e-22;
  Matches   97;  Conservative   66;  Mismatches  121;  Indels   69;  Gaps   12;

Qy         65 NSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSS---------LSTKL 115
              :|||:|       | |           |  |:    :    :|  :         ::
Db          9 DSFEDF-------WKG-----------EDLSNYSYSSTLPPFLLDAAPCEPESLEINKYF 50

Qy        116 IPAIYLLVFVVGVPANA-VTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLN 174
```

```
Db          51 VVIIYALVFLLSLLGNSLVMLVILYSRVGRSVTDVYLLNLALADLLFALTLPIWAASKVN 110

Qy         175 GNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYAL-VTC 233

Db         111 G--WIFGTFLCKVVSLLKEVNFYSGILLLACISVDRYLAIVH--ATRTLTQKRYLVKFIC 166

Qy         234 GLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIP 293

Db         167 LSIWGLSLLLALPVLLFRRTVYSSNVSPACYEDMGN---NTANWRMLLRILPQSFGFIVP 223

Qy         294 FVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNIILII------- 342

Db         224 LLIMLFCYGFTLRTLFKAHMGQKHRAMRVIFA---VVLIFLLCWLPYNLVLLADTLMRTQ 280

Qy         343 ---------HHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNH 386

Db         281 VIQETCERRNHIDRALDATE----------ILGILHSCLNPLIYAFIGQKFRH 323


RESULT 23
S68208
G protein-coupled receptor 12A - human
C;Species: Homo sapiens (man)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 29-Sep-1999
C;Accession: S68208
R;An, S.; Tsai, C.; Goetzl, E.J.
FEBS Lett. 375, 121-124, 1995
A;Title: Cloning, sequencing and tissue distribution of two related G protein-coupled re
A;Reference number: S68207; MUID:96087098; PMID:7498459
A;Accession: S68208
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-365 <ANS>
A;Cross-references: EMBL:U35398; NID:g1015418; PIDN:AAA79060.1; PID:g1015419
C;Superfamily: G protein-coupled receptor 4
C;Keywords: G protein-coupled receptor

  Query Match             16.3%;  Score 351;  DB 2;  Length 365;
  Best Local Similarity   28.8%;  Pred. No. 6.6e-22;
  Matches   90;  Conservative   59;  Mismatches  127;  Indels   36;  Gaps    9;

Qy         104 MGYLT---SSLS--------TKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTT-VFY 151

Db           1 MGNITADNSSMSCTIDHTIHQTLAPVVYVTVLVVGFPANCLSLYFGYLQIKARNELGVYL 60

Qy         152 TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRY 211

Db          61 CNLTVADLFYICSLPFWLQYVLQHDNWSHGDLSCQVCGILLYENIYISVGFLCCISVDRY 120

Qy         212 LAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTC 271

Db         121 LAVAHPFRFHQFRTLKAAVRVTVVIWAKELLTSI--------YFLMHEEVIEDENQHRVC 172

Qy         272 ESSSPFQLY-----YFISLAFFGFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVKAS 322

Db         173 FEHYPIQAWQRAINYYRFLV--GFLFPICLLLASYQGILRAVRRSHGTQKSRKDQIQRLV 230

Qy         323 LLILVIFTICFAPSNIILIIHHANYYYNNTD---GLYFIYLIALCLGSLNSCLDPFLYFL 379

Db         231 LSTVVIFLACFLPYHVLLLVR--SVWEASCDFAKGVFNAYHFSLLLTSFNCVADPVLYCF 288

Qy         380 MSKTRNHSTAYL 391

Db         289 VSETTHRDLARL 300


RESULT 24
JC5653
G protein-coupled receptor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 17-Nov-1997 #sequence_revision 17-Nov-1997 #text_change 21-Jul-2000
C;Accession: JC5653
R;Ferrer-Martinez, A.; Felipe, A.; Mata, J.F.; Casado, F.J.; Pastor-Anglada, M.
Biochem. Biophys. Res. Commun. 238, 107-112, 1997
A;Title: Molecular cloning of a bovine renal G-protein coupled receptor gene (bRGR): Re
A;Reference number: JC5653; MUID:97445134; PMID:9299461
A;Accession: JC5653
A;Molecule type: mRNA
A;Residues: 1-361 <FER>
A;Cross-references: GB:U88366; NID:g2827875; PIDN:AAC05611.1; PID:g2827876
A;Experimental source: renal epithelial cell
C;Superfamily: G protein-coupled receptor 4
C;Keywords: glycoprotein; phosphoprotein
F;24-46/Domain: transmembrane #status predicted <TM1>
F;58-79/Domain: transmembrane #status predicted <TM2>
F;96-117/Domain: transmembrane #status predicted <TM3>
F;137-159/Domain: transmembrane #status predicted <TM4>
F;189-207/Domain: transmembrane #status predicted <TM5>
F;229-253/Domain: transmembrane #status predicted <TM6>
F;265-289/Domain: transmembrane #status predicted <TM7>
F;3,8,62/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;134,218,293,313/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #statu
F;204,221,328/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pr

  Query Match             16.1%;  Score 347;  DB 2;  Length 361;
  Best Local Similarity   27.4%;  Pred. No. 1.4e-21;
  Matches   86;  Conservative   61;  Mismatches  127;  Indels   40;  Gaps    8;

Qy         104 MGYLTS----------SLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTT-VFY 151

Db           1 MGNITADNTSMNCDIDHTIHQTLAPVVYVMVLVVGFPANCLSLYYGYLQIKARNELGVYL 60

Qy         152 TNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRY 211

Db          61 CNLTVADLFYICSLPFWLQYVLQHDHWSHDDLSCQVCGILLYENIYISVGFLCCISIDRY 120

Qy         212 LAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTC 271

Db         121 LAVAHPFRFHQFRTLKAAMGVSALIWVKELLTSI--------YFLMHEEVVEDADRHRVC 172

Qy         272 ESSSPFQ-------LYYFISLAFFGFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVK 320

Db         173 FEHYPLEPRQRGINYYRFL----VGFLFPICLLLASYRGILRAVRRSHGTQKSRKDQIQR 228

Qy         321 ASLLILVIFTICFAPSNIILIIHHANYYYNNTD---GLYFIYLIALCLGSLNSCLDPFLY 377

Db         229 LVLSTVVIFLACFLPYHVLLLVR--SLWESSCDFAKGIFNAYHFSLLLTSFNCVADPVLY 286

Qy         378 FLMSKTRNHSTAYL 391

Db         287 CFVSETTHRDLARL 300


RESULT 25
S63666
platelet activating factor receptor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S63666
R;Ishii, S.; Matsuda, Y.; Nakamura, M.; Waga, I.; Kume, K.; Izumi, T.; Shimizu, T.
Biochem. J. 314, 671-678, 1996
A;Title: A murine platelet-activating factor receptor gene: cloning, chromosomal locali
A;Reference number: S63666; MUID:96239129; PMID:8670084
A;Accession: S63666
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-341 <ISH>
A;Cross-references: EMBL:D50872; NID:g1256924; PIDN:BAA09468.1; PID:g1256925
C;Superfamily: ATP receptor P2u

  Query Match             16.0%;  Score 346;  DB 2;  Length 341;
  Best Local Similarity   29.1%;  Pred. No. 1.6e-21;
  Matches   92;  Conservative   49;  Mismatches  141;  Indels   34;  Gaps   10;

Qy         115 LIPAIYLLVFVVGVPANAVTLWM---LFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAY 171
```

```
Db         17 LFPIVYSVIFILGVVANGYVLWVFANLYPSKKLNEIKIFMVNLTMADLLFLITLPLWIVY 76

Qy        172 HLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALV 231

Db         77 YYNEGDWILPNFLCNVAGCLFFINTYCSVAFLGVITYNRYQAVAYPIKTAQATTRKRGIS 136

Qy        232 TCGLVWATVFLYMLPFFILKQEYYLV-----QPDITTCHDVHNTCESSSPFQLYYFISLA 286

Db        137 LSLIIWVSI-VATASYFLATDSTNLVPNKDGSGNITRCFE-HYEPYSVPILVVHVFIAFC 194

Qy        287 FFGFLIPFVLIIYCYAAIIRTL----------NAYDHRWLWYVKASLLILVIFTICFAPS 336

Db        195 FF---LVFFLIFYCNLVIIHTLLTQPMRQQRKAGVKRRALWMV---CTVLAVFIICFVPH 248

Qy        337 NIILI---IHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNHSTAYLT 392

Db        249 HVVQLPWTLAELGYQTNFHQAINDAHQITLCLLSTNCVLDPVIYCFLTKKFRKH----LS 304

Qy        393 KXXNDLREQGQPSQRT 408

Db        305 EKFYSMRSSRKCSRAT 320
```

RESULT 26
A48921
interleukin-8 receptor type B - mouse
N;Alternate names: G-protein coupled receptor Gpcr16
C;Species: Mus musculus (house mouse)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C;Accession: A48921; A53677; I49348; I55421; H48909; I53774
R;Cerretti, D.P.; Nelson, N.; Kozlosky, C.J.; Morrissey, P.J.; Copeland, N.G.; Gilbert,
Genomics 18, 410-413, 1993
A;Title: The murine homologue of the human interleukin-8 receptor type B maps near the i
A;Reference number: A48921; MUID:94117014; PMID:8288247
A;Accession: A48921
A;Molecule type: DNA
A;Residues: 1-359 <CER>
A;Cross-references: GB:L23637; NID:g435093; PIDN:AAA39305.1; PID:g435094
R;Suzuki, H.; Prado, G.N.; Wilkinson, N.; Navarro, J.
J. Biol. Chem. 269, 18263-18266, 1994
A;Title: The N terminus of interleukin-8 (IL-8) receptor confers high affinity binding t
A;Reference number: A53677; MUID:94308043; PMID:7518426
A;Accession: A53677
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-359 <SUZ>
A;Cross-references: GB:L26549
A;Note: sequence extracted from NCBI backbone (NCBIP:149812)
R;Lee, J.; Cacalano, G.; Camerato, T.; Toy, K.; Moore, M.W.; Wood, W.I.
J. Immunol. 155, 2158-2164, 1995
A;Title: Chemokine binding and activities mediated by the mouse IL-8 receptor.
A;Reference number: I49348; MUID:95363183; PMID:7636264
A;Accession: I49348
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-359 <RES>
A;Cross-references: EMBL:U31207; NID:g950174; PIDN:AAC52239.1; PID:g950175
R;Bozic, C.R.; Gerard, N.P.; von Uexkull-Guldenband, C.; Kolakowski, L.F.
J. Biol. Chem. 269, 29355-29358, 1994
A;Title: The murine interleukin 8 type B receptor homologue and its ligands. Expression
A;Reference number: I55421; MUID:95050766; PMID:7961909
A;Accession: I55421
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-359 <RE2>
A;Cross-references: GB:L13239; NID:g293665; PIDN:AAA62109.1; PID:g293666
R;Wilkie, T.M.; Chen, Y.; Gilbert, D.J.; Moore, K.J.; Yu, L.; Simon, M.I.; Copeland, N.G
Genomics 18, 175-184, 1993
A;Title: Identification, chromosomal location, and genome organization of mammalian G-pr
A;Reference number: A48909; MUID:94116980; PMID:8288218
A;Accession: H48909
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 145-258 <WIL>
A;Cross-references: GB:L20337; NID:g438800; PIDN:AAA16853.1; PID:g438801
R;Harada, A.; Kuno, K.; Nomura, H.; Mukaida, N.; Murakami, S.; Matsushima, K.
Gene 142, 297-300, 1994
A;Title: Cloning of a cDNA encoding a mouse homolog of the interleukin-8 receptor.
A;Reference number: I53774; MUID:94252584; PMID:8194768
A;Accession: I53774
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-359 <RE3>
A;Cross-references: GB:D17630; NID:g493671; PIDN:BAA04536.1; PID:g493672
C;Genetics:
A;Gene: Il8rb
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;49-74/Domain: transmembrane #status predicted <TM1>
F;84-106/Domain: transmembrane #status predicted <TM2>
F;120-141/Domain: transmembrane #status predicted <TM3>
F;163-182/Domain: transmembrane #status predicted <TM4>
F;213-234/Domain: transmembrane #status predicted <TM5>
F;251-271/Domain: transmembrane #status predicted <TM6>
F;308-328/Domain: transmembrane #status predicted <TM7>

Query Match 16.0%; Score 345.5; DB 2; Length 359;
Best Local Similarity 29.8%; Pred. No. 1.9e-21;
Matches 96; Conservative 59; Mismatches 120; Indels 47; Gaps 11;

```
Qy        111 LSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC--TTVFYTNLAIADFLFCVTLPFK 168

Db         45 INSYAVVVIYVLVTLLSLVGNSLVMLVILY-NRSTCSVTDVYLLNLAIADLFFALTLPVW 103

Qy        169 IAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTY 228

Db        104 AASKVNG--WTFGSTLCKIFSYVKEVTFYSSVLLLACISMDRYLAIVHATSTLIQKRHLV 161

Qy        229 ALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFF 288

Db        162 KFV-CIAMWLLSVILALPILILRNPVKVNLSTLVCYEDVGN---NTSRLRVVLRILPQTF 217

Qy        289 GFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNIILIIHH 344

Db        218 GFLVPLLIMLFCYGFTLRTLFKAHMGQKHRAMRVIFA---VVLVFLLCWLPYNLVLF--- 271

Qy        345 ANYYYNNTDGLYFIYLI-----------AL----CLGSLNSCLDPFLYFLMSKTRNHS- 387

Db        272 -------TDTLMRTKLIKETCERRDDIDKALNATEILGFLHSCLNPIIYAFIGQKFRHGL 324

Qy        388 ----TAYLTKXXNDLREQGQPS 405

Db        325 LKIMATYGLVSKEFLAKEGRPS 346
```

RESULT 27
B57641
G protein-coupled receptor 4 - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 29-Sep-1999
C;Accession: B57641
R;Mahadevan, M.S.; Baird, S.; Bailly, J.E.; Shutler, G.G.; Sabourin, L.A.; Tsilfidis, C
Genomics 30, 84-88, 1995
A;Title: Isolation of a novel G protein-coupled receptor (GPR4) localized to chromosome
A;Reference number: A57641; MUID:96129306; PMID:8595909
A;Accession: B57641
A;Status: preliminary; nucleic acid sequence not shown; translation not shown; signific
A;Molecule type: DNA
A;Residues: 1-362 <MAH>
A;Cross-references: GB:U22108; NID:g722282; PIDN:AAA98458.1; PID:g722283
C;Genetics:
A;Introns: #status absent
C;Superfamily: G protein-coupled receptor 4
C;Keywords: G protein-coupled receptor

```
  Query Match             16.0%;  Score 345.5;  DB 2;  Length 362;
  Best Local Similarity   29.7%;  Pred. No. 1.9e-21;
  Matches   84;  Conservative   53;  Mismatches  117;  Indels   29;  Gaps    7;

Qy        117 PAIYLLVFVVGVPANAVTLWMLFFRTRSICTT-VFYTNLAIADFLFCVTLPFKIAYHLNG 175
              |::|: |  ||:| | : ||  : : |      |:  ||:||| |:  |||  : | |:
Db         21 PSLYIFVIGVGLPTNCLRLWAAYRQVRQRNELGVYLMNLSIADLLYICTLPLWVDYFLHH 80

Qy        176 NNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGL 235
              :||: |   |:    ||| |:| ||  | |||::|||: ||  :  | :   |:    :
Db         81 DNWIHGPGSCKLFGFIFYTNIYISIAFLCCISVDRYLAVAHPLRFARLRRVKTAVAVSSV 140

Qy        236 VWAT-VFLYMLPFF---ILKQEYYLVQPDITTCHDVHNTCESSSPFQ---LYYFISLAFF 288
              |||| :    :| |   : :  |             | |    | :    :  :   |
Db        141 VWATELGANSVPLFHDELFRDRYN-----------HTFCFEKFPMEGWVAWMNLYRVFV 188

Qy        289 GFLIPFVLIIYCYAAIIR----TLNAYDHRWLWYVKASLLILVIFTICFAPSNIILIIHH 344
              ||| |: |::  |  |:|    :::          : :| :: |  :|||| :::|:
Db        189 GFLFPWALMLLSYRGILRAVRGSVSTERQEKAKIKRLALSLIAIVLVCFAPYHVLLLSRS 248

Qy        345 ANYYYNNTD-----GLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
              | |  :  |      ::  |  :|   |||   || || |:::
Db        249 AVYLGHPWDCGFEERVFSAYHSSLAFTSLNCVADPILYCLVNE 291


RESULT 28
S15403
angiotensin II receptor type 1 - bovine
C;Species: Bos primigenius indicus x Bos primigenius taurus (cattle)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C;Accession: S15403
R;Sasaki, K.; Yamano, Y.; Bardhan, S.; Iwai, N.; Murray, J.J.; Hasegawa, M.; Matsuda, Y.
Nature 351, 230-233, 1991
A;Title: Cloning and expression of a complementary DNA encoding a bovine adrenal angiote
A;Reference number: S15403; MUID:91251900; PMID:2041569
A;Accession: S15403
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-359 <SAS>
A;Cross-references: GB:X62294; NID:g43; PIDN:CAA44182.1; PID:g44
C;Superfamily: vertebrate rhodopsin

  Query Match             16.0%;  Score 345;  DB 2;  Length 359;
  Best Local Similarity   30.7%;  Pred. No. 2.1e-21;
  Matches   89;  Conservative   47;  Mismatches  120;  Indels   34;  Gaps    6;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTR-SICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              :|| :| ::||||:  |:: : :::|  :     :||  |||:||  | :|||    |
Db         30 MIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLTLPLWAVYTA 89

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
                | ||  ||:  :     |:| |: || |:||:|||||||  |       | |||
Db         90 MEYRWPFGNYLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTC 149

Qy        234 GLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIP 293
              ::|    |  ||  | :  :::   :|| |   : :  |: |  |   ::     ||| |
Db        150 IIIWLLAGLASLPTIIHRNVFFIENTNITVCAFHYESQNSTLPVGL--GLTKNILGFLFP 207

Qy        294 FVLIIYCYAAIIRTL-NAYD-----HRWLWYVKASLLILVIFTICFAPSNIILIIHHANY 347
              |::|:  |  | :||  ||:      |     |  | |:: |   : |  |   :
Db        208 FLIILTSYTLIWKTLKKAYEIQKNKPRKDDIFKIILAIVLFFFFSWVPHQIFTFM----- 262

Qy        348 YYNNTDGLYFIYL--------------IALCLGSLNSCLDPFLYFLMSK 382
                  | |  : |               | :||   |:||:|  |  : |
Db        263 -----DVLIQLGLIRDCKIEDIVDTAMPITICLAYFNNCLNPLFYGFLGK 307


RESULT 29
S26667
G protein-coupled receptor BLR1 - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C;Accession: S26667
R;Dobner, T.; Wolf, I.; Emrich, T.; Lipp, M.
Eur. J. Immunol. 22, 2795-2799, 1992
A;Title: Differentiation-specific expression of a novel G protein-coupled receptor from
A;Reference number: S26667; MUID:93049615; PMID:1425907
A;Accession: S26667
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-372 <DOB>
A;Cross-references: EMBL:X68149; NID:g29459; PIDN:CAA48252.1; PID:g29460
C;Genetics:
A;Gene: GDB:BLR1
A;Cross-references: GDB:136235; OMIM:601613
A;Map position: 15q26.1-15q26.1
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             15.9%;  Score 344;  DB 2;  Length 372;
  Best Local Similarity   33.0%;  Pred. No. 2.6e-21;
  Matches  104;  Conservative   51;  Mismatches  130;  Indels   30;  Gaps   15;

Qy         98 HVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWML--FFRTRSICTTVFYTNLA 155
              |:  || | | :|     :|  | |:|::||  | : | :|    :|||  |  |  :||
Db         36 HLCPATEGPLMASFKAVFVPVAYSLIFLLGVIGNVLVLVILERHRQTRS-STETFLFHLA 94

Qy        156 IADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIV 215
              :|| |    ||| :|   |  || |  ||:    :  | ||| ||||||:::||||||
Db         95 VADLLLVFILPFAVAEGSVG--WVLGTFLCKTVIALHKVNFYCSSLLLACIAVDRYLAIV 152

Qy        216 HP-FTYRGLPKHTYAL---VTCGLVWATVFLYMLP---FFILKQEYYLVQPDITTC-HDV 267
              |    ||    |   |  :||| :|   ||  ||   |  : | ::     :  |
Db        153 HAVHAYR----HRRLLSIHITCGTIWLVGFLLALPEILFAKVSQGHH--NNSLPRCTFSQ 206

Qy        268 HNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHR--WLWYVKASLLI 325
              |  |: : |   :   :|  |||:| ::: :||  ::  |.     |       |: ::|:
Db        207 ENQAETHAWFTSRFLYHVA--GFLLPMLVMGWCYVGVVHRLRQAQRRPQRQKAVRVAILV 264

Qy        326 LVIFTICFAPSNIILIIHHANYY--YNNTDGL--YFIYLIALC--LGSLNSCLDPFLY-F 378
              || :|::| :|:: :          :||  |       | :|  ||  : ||:| || |
Db        265 TSIFFLCWSPYHIVIFLDTLARLKAVDNTCKLNGSLPVAITMCEFLGLAHCCLNPMLYTF 324

Qy        379 LMSKTRNHSTAYLTK 393
                | |:  :  |||
Db        325 AGVKFRSDLSRLLTK 339


RESULT 30
I51372
angiotensin II receptor - turkey
C;Species: Meleagris gallopavo (common turkey)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Nov-1999
C;Accession: I51372; PN0449
R;Murphy, T.J.; Nakamura, Y.; Takeuchi, K.; Alexander, R.W.
Mol. Pharmacol. 44, 1-7, 1993
A;Title: A cloned angiotensin receptor isoform from the turkey adrenal gland is pharmac
A;Reference number: I51372; MUID:93341466; PMID:8341266
A;Accession: I51372
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-359 <MUR>
A;Cross-references: GB:L23203; NID:g349735; PIDN:AAA03560.1; PID:g349736
R;Carsia, R.V.; McIlroy, P.J.; Kowalski, K.I.; Tilly, J.L.
Biochem. Biophys. Res. Commun. 191, 1073-1080, 1993
A;Title: Isolation of turkey adrenocortical cell angiotensin II (AII) receptor partial
A;Reference number: PN0449; MUID:93221469; PMID:7916599
A;Accession: PN0449
A;Molecule type: mRNA
A;Residues: 53-91,'C',93-162,'SSFIVIY',171-291 <CAR>
A;Cross-references: GB:S58041; NID:g299373; PIDN:AAB26041.1; PID:g299374
A;Experimental source: liver
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; phosphoprotein; transmembrane protein
F;53-64/Domain: intracellular #status predicted <IN1>
```

```
F;65-86/Domain: transmembrane #status predicted <TM1>
F;87-102/Domain: extracellular #status predicted <EX1>
F;103-124/Domain: transmembrane #status predicted <TM2>
F;125-142/Domain: intracellular #status predicted <IN2>
F;143-162/Domain: transmembrane #status predicted <TM3>
F;163-192/Domain: extracellular #status predicted <EX2>
F;193-214/Domain: transmembrane #status predicted <TM4>
F;215-238/Domain: intracellular #status predicted <IN3>
F;239-262/Domain: transmembrane #status predicted <TM5>
F;263-274/Domain: extracellular #status predicted <EX3>
F;275-291/Domain: transmembrane #status predicted <TM6>
F;136/Binding site: phosphate (Ser) (covalent) #status predicted
F;141,233/Binding site: phosphate (Thr) (covalent) #status predicted
F;176,188/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             15.9%;  Score 343;  DB 2;  Length 359;
  Best Local Similarity   27.8%;  Pred. No. 3e-21;
  Matches   88;  Conservative   57;  Mismatches  130;  Indels   42;  Gaps    8;

Qy          91 EESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLF-FRTRSICTTV 149
               ||:   :||     |    |    ::| :| ::|::|:  |:: : ::: :       ::
Db           8 EETVKRIHVDCPVSG--RHSYIYIMVPTVYSIIFIIGIFGNSLVVIVIYCYMKLKTVASI 65

Qy         150 FYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISIN 209
               |  |||:||  | :|||   ||      | ||  ||:  :     |:| |: || |:||:
Db          66 FLLNLALADLCFLITLPLWAAYTAMEYQWPFGNCLCKLASAGISFNLYASVFLLTCLSID 125

Qy         210 RYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTC---HD 266
               ||||||||   |       | ||| ::|    :  ||  | :  ::    ::| |   :|
Db         126 RYLAIVHPVKSRIRRTMFVARVTCIVIWLLAGVASLPVIIHRNIFFAENLNMTVCGFRYD 185

Qy         267 VHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWYVKA----- 321
                :||       ::   :|    ||||||::|:  |  | :||           ||
Db         186 NNNT-----TLRVGLGLSKNLLGFLIPFLIILTSYTLIWKTLK----------KAYQIQR 230

Qy         322 ---------SLLILVIFTICFA--PSNI-----ILIIHHANYYYNNTDGLYFIYLIALCL 365
                      ::: ::|   |:  |  :     :||  |        || :       :|:
Db         231 NKTRNDDIFKMIVAIVFFFFFSWIPHQVFTFLDVLIQLHVITDCKITDIVDTAMPFTICI 290

Qy         366 GSLNSCLDPFLYFLMSK 382
                 |:||:|| |    |
Db         291 AYFNNCLNPFFYVFFGK 307


RESULT 31
A40191
platelet-activating factor receptor - human
C;Species: Homo sapiens (man)
C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 20-Jun-2000
C;Accession: A40191; JH0479; A41079; JC1359; A42831; I51923
R;Kunz, D.; Gerard, N.P.; Gerard, C.
J. Biol. Chem. 267, 9101-9106, 1992
A;Title: The human leukocyte platelet-activating factor receptor. cDNA cloning, cell sur
A;Reference number: A40191; MUID:92250505; PMID:1374385
A;Accession: A40191
A;Molecule type: mRNA
A;Residues: 1-342 <KUN>
A;Cross-references: GB:M76674; NID:g456293; PIDN:AAA60002.1; PID:g456294
R;Ye, R.D.; Prossnitz, E.R.; Zou, A.; Cochrane, C.G.
Biochem. Biophys. Res. Commun. 180, 105-111, 1991
A;Title: Characterization of a human cDNA that encodes a functional receptor for platele
A;Reference number: JH0479; MUID:92028922; PMID:1656963
A;Accession: JH0479
A;Molecule type: mRNA
A;Residues: 1-342 <YER>
A;Cross-references: GB:M80436; NID:g189537; PIDN:AAA60001.1; PID:g189538
A;Experimental source: granulocyte, cell line HL-60 all
R;Nakamura, M.; Honda, Z.; Izumi, T.; Sakanaka, C.; Mutoh, H.; Minami, M.; Bito, H.; Sey
J. Biol. Chem. 266, 20400-20405, 1991
A;Title: Molecular cloning and expression of platelet-activating factor receptor from hu
A;Reference number: A41079; MUID:92041873; PMID:1657923
A;Accession: A41079
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-342 <NAK>
A;Cross-references: GB:D10202; GB:D90433; NID:g219975; PIDN:BAA01050.1; PID:g219976
R;Sugimoto, T.; Tsuchimochi, H.; McGregor, C.G.A.; Mutoh, H.; Shimizu, T.; Kurachi, Y.
Biochem. Biophys. Res. Commun. 189, 617-624, 1992
A;Title: Molecular cloning and characterization of the platelet-activating factor recep
A;Reference number: JC1359; MUID:93112021; PMID:1281995
A;Accession: JC1359
A;Molecule type: mRNA
A;Residues: 1-315,'N',317-342 <SUG>
A;Experimental source: heart
A;Note: the authors translated the codon AAT for residue 316 as Lys
R;Seyfried, C.E.; Schweickart, V.L.; Godiska, R.; Gray, P.W.
Genomics 13, 832-834, 1992
A;Title: The human platelet-activating factor receptor gene (PTAFR) contains no introns
A;Reference number: A42831; MUID:92347886; PMID:1322356
A;Accession: A42831
A;Molecule type: DNA
A;Residues: 1-226,'TG',229-342 <SEY>
A;Cross-references: GB:M88177; NID:g190697; PIDN:AAA60214.1; PID:g190698
A;Note: sequence extracted from NCBI backbone (NCBIN:109813, NCBIP:109814)
R;Chase, P.B.; Halonen, M.; Regan, J.W.
Am. J. Respir. Cell Mol. Biol. 8, 240-244, 1993
A;Title: Cloning of a human platelet-activating factor receptor gene: evidence for an i
A;Reference number: I51923; MUID:93192035; PMID:8383507
A;Accession: I51923
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-342 <RES>
A;Cross-references: GB:S56396; NID:g298580; PIDN:AAB25755.1; PID:g298581
C;Genetics:
A;Gene: GDB:PTAFR
A;Cross-references: GDB:128806; OMIM:173393
A;Map position: 1p35-1p34.3
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein
F;17-38/Domain: transmembrane #status predicted <TRI>
F;54-75/Domain: transmembrane #status predicted <TII>
F;92-113/Domain: transmembrane #status predicted <III>
F;134-155/Domain: transmembrane #status predicted <TIV>
F;184-205/Domain: transmembrane #status predicted <TRV>
F;233-253/Domain: transmembrane #status predicted <TVI>
F;277-297/Domain: transmembrane #status predicted <VII>

  Query Match             15.9%;  Score 342.5;  DB 2;  Length 342;
  Best Local Similarity   28.3%;  Pred. No. 3.2e-21;
  Matches   96;  Conservative   50;  Mismatches  132;  Indels   61;  Gaps   14;

Qy         106 YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-----TTVFYTNLAIADFL 160
               :: |      | | :| ::||:|| ||   ||:  |      |       :|  || :|| |
Db           8 HMDSEFRYTLFPIVYSIIFVLGVIANGYVLWV--FARLYPCKKFNEIKIFMVNLTMADML 65

Qy         161 FCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTY 220
               | :|||  | |: |  ||:  : ||     :|: | |||:  |  |: ||: |:  |
Db          66 FLITLPLWIVYYQNQGNWILPKFLCNVAGCLFFINTYCSVAFLGVITYNRFQAVTRPIKT 125

Qy         221 RGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLV------QPD------ITTCHDVH 268
                      :   ::|  :       :    |:|:       ||       :| | : |
Db         126 AQANTRKRGISLSLVIWVAI--------VGAASYFLILDSTNTVPDSAGSGNVTRCFE-H 176

Qy         269 NTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTL---------NA-YDHRWLWY 318
                  |     :: ||    | | : |::|::|   |||||         ||    | ||
Db         177 YEKGSVPVLIIHIFI---VFSFFLVFLIILFCNLVIIRTLLMQPVQQQRNAEVKRRALWM 233

Qy         319 VKASLLILVIFTICFAPSNII--------LIIHHANYYYNNTDGLYFIYLIALCLGSLNS 370
               |     :| :| ||| | :::        |    : ::    |     : : ||| | |
Db         234 V---CTVLAVFIICFVPHHVVQLPWTLAELGFQDSKFHQAINDA----HQVTLCLLSTNC 286

Qy         371 CLDPFLY-FLMSKTRNHSTAYLTKXXNDLREQGQPSQRT 408
                ||| :| ||  | | |    ||:    :|   : |: |
Db         287 VLDPVIYCFLTKKFRKH----LTEKFYSMRSSRKCSRAT 321
```

```
RESULT 32
A53752
interleukin-8 receptor (clone 5B1a) - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C;Accession: A53752
R;Prado, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.; Wilkinson, N.; Folco, E.; Navarro
J. Biol. Chem. 269, 12391-12394, 1994
A;Title: Molecular characterization of a novel rabbit interleukin-8 receptor isotype.
A;Reference number: A53752; MUID:94230294; PMID:8175642
A;Accession: A53752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-358 <PRA>
A;Cross-references: GB:L24445; NID:g437661; PIDN:AAA31378.1; PID:g437662
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             15.9%;  Score 342.5;  DB 2;  Length 358;
  Best Local Similarity   30.8%;  Pred. No. 3.3e-21;
  Matches   93;  Conservative   59;  Mismatches  105;  Indels   45;  Gaps   12;

Qy        108 TSSLSTK--LIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC--TTVFYTNLAIADFLFCV 163
              : || |   ::   |:|||:: :  |:: : :: : :|| |  | |:  |||||| ||
Db         39 SESLETNSYVVLITYILVFLLSLLGNSLVMLVILY-SRSTCSVTDVYLLNLAIADLLFAT 97

Qy        164 TLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGL 223
              |||   |  ::|  | ||  ||:  :::   | |  ||||||||::|||||||  |   :
Db         98 TLPIWAASKVHG--WTFGTPLCKVVSLVKEVNFYSGILLLACISVDRYLAIVHA-TRTMI 154

Qy        224 PKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFI 283
              |       |  :|    :  ||  : :   :           |: |   |:: :::   |
Db        155 QKRHLVKFICLSMWGVSLILSLPILLFRNAIFPPNSSPVCYEDMGN---STAKWRMVLRI 211

Qy        284 SLAFFGFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNII 339
                 |||::| :::::||   :|||        || :  : |   :::|| :|: | |::
Db        212 LPQTFGFILPLLVMLFCYVFTLRTLFQAHMGQKHRAMRVIFA---VVLIFLLCWLPYNLV 268

Qy        340 LIIHHANYYYNNTDGLYFIYLI-----------AL----CLGSLNSCLDPFLY-FLMSK 382
              |:          || |   ::|           ||     || |:|||:| :| |:  |
Db        269 LL----------TDTLMRTHVIQETCERRNDIDRALDATEILGFLHSCLNPIIYAFIGQK 318

Qy        383 TR 384
              |
Db        319 FR 320


RESULT 33
S32785
G protein-coupled receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C;Accession: S32785
R;Kouba, M.; Vanetti, M.; Wang, X.; Schaefer, M.; Hoellt, V.
FEBS Lett. 321, 173-178, 1993
A;Title: Cloning of a novel putative G-protein-coupled receptor (NLR) which is expressed
A;Reference number: S32785; MUID:93238948; PMID:8386678
A;Accession: S32785
A;Molecule type: mRNA
A;Residues: 1-374 <KOU>
A;Cross-references: GB:X71463; GB:S59748; NID:g599926; PIDN:CAA50582.1; PID:g599927
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             15.9%;  Score 342;  DB 2;  Length 374;
  Best Local Similarity   31.2%;  Pred. No. 3.8e-21;
  Matches  102;  Conservative   50;  Mismatches  107;  Indels   68;  Gaps   16;

Qy        105 GYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWML--FFRTRSICTTVFYTNLAIADFLFC 162
              | | :|  |  :|  | |:|::|:  | : | :|     |||  |  |  :||:|| |
Db         45 GPLLTSFKTIFMPVAYSLIFLLGMMGNILVLVILERHRHTRS-STETFLFHLAVADLLLV 103

Qy        163 VTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRG 222
              ||| :|    |  || |  ||:    :   | ||| ||||||:::|||||||
Db        104 FILPFAVAEGSVG--WVLGTFLCKTVIALHKINFYCSSLLLACIAVDRYLAIVHAV---- 157

Qy        223 LPKHTYAL-------VTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSS 275
              | |          :||  :|   ||: ||  :  :   :|||         ||    |
Db        158 ---HAYRRRRLLSIHITCSTIWLAGFLFALPELLFAK---VVQP--------HN--NESL 201

Qy        276 PFQLY----------YFISLAFF---GFLIPFVLIIYCYAAIIRTLNAYDHRWLW----- 317
              |  ::          :| |   :   |||:| ::: :||  ::        || |
Db        202 PQCIFSQENEAETRAWFASRFLYHTGGFLLPMLVMAWCYVGVV-------HRLLQAQRRP 254

Qy        318 ----YVKASLLILVIFTICFAPSNIILIIHHANYY--YNNTDGL--YFIYLIALC--LGS 367
                  |: ::|:  || :|::| :|:: :          |::  |  |    | ||  ||
Db        255 QRQKAVRVAILVTSIFLLCWSPYHIVIFLDTLERLKAVNSSCELSGYLSVAITLCEFLGL 314

Qy        368 LNSCLDPFLY-FLMSKTRNHSTAYLTK 393
               : ||:|. || |   | |:  :  |||
Db        315 AHCCLNPMLYTFAGVKFRSDLSRLLTK 341


RESULT 34
A48227
kappa opioid receptor 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jan-2000
C;Accession: A48227; JC4138
R;Yasuda, K.; Raynor, K.; Kong, H.; Breder, C.D.; Takeda, J.; Reisine, T.; Bell, G.I.
Proc. Natl. Acad. Sci. U.S.A. 90, 6736-6740, 1993
A;Title: Cloning and functional comparison of kappa and delta opioid receptors from mou
A;Reference number: A48227; MUID:93342064; PMID:8393575
A;Accession: A48227
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-380 <YAS>
A;Cross-references: GB:L11065; NID:g348248; PIDN:AAA39363.1; PID:g348249
R;Liu, H.C.; Lu, S.; Augustin, L.B.; Felsheim, R.F.; Chen, H.C.; Loh, H.H.; Wei, L.N.
Biochem. Biophys. Res. Commun. 209, 639-647, 1995
A;Title: Cloning and promoter mapping of mouse kappa opioid receptor gene.
A;Reference number: JC4138; MUID:95251663; PMID:7733933
A;Accession: JC4138
A;Molecule type: mRNA
A;Residues: 1-380 <LIU>
A;Note: The authors translated the codon CAG for residue 365 as Glu
C;Comment: This receptor exists in different areas of the central and peripheral nervou
C;Genetics:
A;Gene: kor
C;Superfamily: vertebrate rhodopsin
C;Keywords: brain; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembr

  Query Match             15.9%;  Score 342;  DB 2;  Length 380;
  Best Local Similarity   26.6%;  Pred. No. 3.9e-21;
  Matches   93;  Conservative   62;  Mismatches  156;  Indels   38;  Gaps    8;

Qy         55 PIKTFRGAP----------PNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNAT 103
              ||: ||| |          |||   ||      | :     :   ::     |: |
Db          4 PIQIFRGDPGPTCSPSACLLPNSSSWFP-----NWAESDSNGSVGSEDQQLESAHISPAI 58

Qy        104 MGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTR-SICTTVFYTNLAIADFLFC 162
                      :| |:| :|||||:  |:: ::::   |:   | ::  |||:|| |
Db         59 ---------PVIITAVYSVVFVVGLVGNSLVMFVIIRYTKMKTATNIYIFNLALADALVT 109

Qy        163 VTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRG 222
              |:||: | :|  |:| ||:|||:    | | ||: ||  |  :|::||:|: ||
Db        110 TTMPFQSAVYLM-NSWPFGDVLCKIVISIDYYNMFTSIFTLTMMSVDRYIAVCHPVKALD 168

Qy        223 LPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYF 282
                  | :   :|       :   :|         |:  |  :    :  | : |:
Db        169 FRTPLKAKIINICIWLLASSVGISAIVLGGTKVREDVDVIEC-SLQFPDDEYSWWDLFMK 227

Qy        283 ISLAFFGFLIPFVLIIYCYAAIIRTLNAY---------DHRWLWYVKASLLILVIFTICF 333
              | :  | |:|| ::|| ||  :|  | :         |       |  |::: :| ||:
Db        228 ICVFVFAFVIPVLIIIVCYTLMILRLKSVRLLSGSREKDRNLRRITKLVLVVVAVFIICW 287
```

```
Qy        334 APSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
              | :| :::       ::|  |    |   : ||  || |:| ||  : :
Db        288 TPIHIFILVEALGSTSHSTAAL-SSYYFCIALGYTNSSLNPVLYAFLDE 335


RESULT 35
S13638
platelet-activating factor receptor - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 08-Oct-1999
C;Accession: S13638
R;Honda, Z.; Nakamura, M.; Miki, I.; Minami, M.; Watanabe, T.; Seyama, Y.; Okado, H.; To
Nature 349, 342-346, 1991
A;Title: Cloning by functional expression of platelet-activating factor receptor from gu
A;Reference number: S13638; MUID:91101726; PMID:1846231
A;Accession: S13638
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-342 <HON>
A;Cross-references: GB:X56736; NID:g49442; PIDN:CAA40060.1; PID:g49443
A;Note: the species of guinea pig is not identified; in GenBank entry CCPAFREC, release
C;Superfamily: ATP receptor P2u

  Query Match             15.8%;  Score 341.5;  DB 2;  Length 342;
  Best Local Similarity   28.1%;  Pred. No. 3.9e-21;
  Matches   91;  Conservative   53;  Mismatches  131;  Indels   49;  Gaps   11;

Qy        115 LIPAIYLLVFVVGVPANAVTLWM---LFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAY 171
              | | :| ::||:|: ||   ||:   |:   :     :|  || :|| || :||| | |
Db         17 LFPIVYSIIFVLGIIANGYVLWVFARLYPSKKLNEIKIFMVNLTVADLLFLITLPLWIVY  76

Qy        172 HLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALV 231
              : |  ||   : ||     :|: | |||:  |  |: ||: |: :|            :
Db         77 YSNQGNWFLPKFLCNLAGCLFFINTYCSVAFLGVITYNRFQAVKYPIKTAQATTRKRGIA 136

Qy        232 TCGLVWATVFLYMLPFFILKQEYYLVQP-----------DITTCHDVHNTCESSSPFQL 279
              ::|  :       :    |:||              :|| | : :   : | |  :
Db        137 LSLVIWVAI--------VAAASYFLVMDSTNVVSNKAGSGNITRCFEHYE--KGSKPVLI 186

Qy        280 YYFISLAFFGFLIPFVLIIYCYAAIIRTL---------NA-YDHRWLWYVKASLLILVIF 329
              :  :   || | |:||::|   || ||          ||   | || |     :| :|
Db        187 IHICIV--LGFFIVFLLILFCNLVIIHTLLRQPVKQQRNAEVRRRALWMV---CTVLAVF 241

Qy        330 TICFAPSNII----LIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTR 384
              |||  | :::     :     :  :|   :   : : ||| | |  ||| :| || | |
Db        242 VICFVPHHMVQLPWTLAELGMWPSSNHQAINDAHQVTLCLLSTNCVLDPVIYCFLTKKFR 301

Qy        385 NHSTAYLTKXXNDLREQGQPSQRT 408
              |    |::  | :|   : |: |
Db        302 KH----LSEKLNIMRSSQKCSRVT 321


RESULT 36
JC5068
G protein-coupled receptor CKR-L3 - human
C;Species: Homo sapiens (man)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 21-Jul-2000
C;Accession: JC5068
R;Zaballos, A.; Varona, R.; Gutierrez, J.; Lind, P.; Marquez, G.
Biochem. Biophys. Res. Commun. 227, 846-853, 1996
A;Title: Molecular cloning and RNA expression of two new human chemokine receptor-like g
A;Reference number: JC5067; MUID:97040707; PMID:8886020
A;Accession: JC5068
A;Molecule type: DNA
A;Residues: 1-369 <ZAB>
A;Cross-references: EMBL:Z79784; NID:g1668737; PIDN:CAB02144.1; PID:g1668738
C;Comment: This protein belongs to the family of alpha chemokine receptors.
C;Genetics:
A;Gene: GDB:CMKBR6; STRL22; GPR29; CCR6; CKR-L3; GPR-CY4
A;Cross-references: GDB:5370639; OMIM:601835
A;Map position: 6q27-6q27
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein
F;42-68/Domain: transmembrane #status predicted <TM1>
F;79-99/Domain: transmembrane #status predicted <TM2>
F;115-136/Domain: transmembrane #status predicted <TM3>
F;160-180/Domain: transmembrane #status predicted <TM4>
F;212-233/Domain: transmembrane #status predicted <TM5>
F;250-271/Domain: transmembrane #status predicted <TM6>
F;292-315/Domain: transmembrane #status predicted <TM7>

  Query Match             15.8%;  Score 341.5;  DB 2;  Length 369;
  Best Local Similarity   30.0%;  Pred. No. 4.2e-21;
  Matches   91;  Conservative   50;  Mismatches  133;  Indels   29;  Gaps   10;

Qy        112 STKLIPAIYLLVFVVGVPAN-AVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIA 170
              |   :|  | |: | |:  |  | :   |::     | |:  |:|||| || :||||
Db         40 SRLFVPIAYSLICVFGLLGNILVVITFAFYKKARSMTDVYLLNMAIADILFVLTLPFWAV 99

Qy        171 YHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHP-----FTYRGLPK 225
              |  |  |||     |:    |:  |  | :||| |||::||:|||          | ||:
Db        100 SHATG-AWVFSNATCKLLKGIYAINFNCGMLLLTCISMDRYIAIVQATKSFRLRSRTLPR 158

Qy        226 HTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISL 285
              : : | :||    :     |:  |:|     |:  |   : |       ::|
Db        159 ---SKIICLVVWGLSVIISSSTFVFNQKYNTQGSDV--CEPKYQTVSEPIRWKLLMLGLE 213

Qy        286 AFFGFLIPFVLIIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTICFAPSNIILI 341
              ||| || : :|:||  |::||     |:  |:    ::  : ::::|  |  | |::|:
Db        214 LLFGFFIPLMFMIFCYTFIVKTLVQAQNSKRHK---AIRVIIAVVLVFLACQIPHNMVLL 270

Qy        342 IHHANYYYNNTDG-----LYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNHSTAYLTKXX 395
              :  ||    |        : :   :   |  |: ||:| || |:  | ||    |  |
Db        271 VTAANLGKMNRSCQSEKLIGYTKTVTEVLAFLHCCLNPVLYAFIGQKFRN----YFLKIL 326

Qy        396 NDL 398
              ||
Db        327 KDL 329


RESULT 37
A48857
angiotensin II receptor type 1 - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C;Accession: A48857
R;Burns, K.D.; Inagami, T.; Harris, R.C.
Am. J. Physiol. 264, F645-F654, 1993
A;Title: Cloning of a rabbit kidney cortex AT1 angiotensin II receptor that is present
A;Reference number: A48857; MUID:93236091; PMID:7916579
A;Accession: A48857
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-359 <BUR>
A;Cross-references: GB:S59041; NID:g299614; PIDN:AAB26239.1; PID:g299615
A;Experimental source: proximal tubule cells
A;Note: sequence extracted from NCBI backbone (NCBIN:129600, NCBIP:129601)
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             15.8%;  Score 341;  DB 2;  Length 359;
  Best Local Similarity   30.3%;  Pred. No. 4.5e-21;
  Matches   86;  Conservative   49;  Mismatches  127;  Indels   22;  Gaps    6;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTR-SICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              :|| :| ::||||:  |:: : :::|  :     :||  |||:||  | :|||    |
Db         30 MIPTLYSIIFVVGIFGNSLAVIVIYFYMKLKTVASVFLLNLALADLCFLLTLPLWAVYTA 89

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
              | ||  ||:  :     |:| |: || |:||:||||||||   |       | |||
Db         90 MEYRWPFGNYLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTC 149

Qy        234 GLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIP 293
              ::|    |  ||  | :  :::   :|| |   : :  |: | |   ::    ||| |
Db        150 IIIWLLAGLASLPAIIHRNVFFIENTNITVCAFHYESQNSTLPIGL--GLTKNILGFLFP 207
```

```
Qy        294 FVLIIYCYAAIIRTL-NAYD-----HRWLWYVKASLLILVIFTICFAPSNIIL------- 340
              |::|:  |  | : |  ||:      |    |  : |:: |   :  | |
Db        208 FLIILTSYTLIWKALKKAYEIQKNKPRNDDIFKIIMAIVLFFFFSWVPHQIFTFLDVLIQ 267

Qy        341 --IIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
                :||         | :     | :|:   |:||:|  |  : |
Db        268 LGVIHDCRI----ADIVDTAMPITICIAYFNNCLNPLFYGFLGK 307


RESULT 38
JC2134
angiotensin II receptor type 1A - rat
N;Alternate names: AT1a receptor; AT3 receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 24-Nov-1999
C;Accession: JC2134; S15404; S20424; JQ1055
R;Conchon, S.; Monnot, C.; Sirieix, M.E.; Bihoreau, C.; Corvol, P.; Clauser, E.
Biochem. Biophys. Res. Commun. 199, 1347-1354, 1994
A;Title: Synthetic cDNA encoding the rat AT1a receptor: a useful tool for structure-func
A;Reference number: JC2134; MUID:94197726; PMID:8147879
A;Accession: JC2134
A;Molecule type: mRNA
A;Residues: 1-359 <CON>
A;Note: the amino acid sequence of this protein is not given
R;Murphy, T.J.; Alexander, R.W.; Griendling, K.K.; Runge, M.S.; Bernstein, K.E.
Nature 351, 233-236, 1991
A;Title: Isolation of a cDNA encoding the vascular type-1 angiotensin II receptor.
A;Reference number: S15404; MUID:91251901; PMID:2041570
A;Accession: S15404
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-359 <MUR>
A;Cross-references: GB:X62295; NID:g57773; PIDN:CAA44183.1; PID:g57774
R;Iwai, N.; Inagami, T.
FEBS Lett. 298, 257-260, 1992
A;Title: Identification of two subtypes in the rat type I angiotensin II receptor.
A;Reference number: S20423; MUID:92183879; PMID:1544458
A;Accession: S20424
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-359 <IWA>
R;Iwai, N.; Yamano, Y.; Chaki, S.; Konishi, F.; Bardhan, S.; Tibbetts, C.; Sasaki, K.; H
Biochem. Biophys. Res. Commun. 177, 299-304, 1991
A;Title: Rat angiotensin II receptor: cDNA sequence and regulation of the gene expressio
A;Reference number: JQ1055; MUID:91254291; PMID:2043116
A;Accession: JQ1055
A;Molecule type: mRNA
A;Residues: 1-80,'C',82-108,'T',110-359 <IW2>
A;Cross-references: GB:M74054; NID:g202918; PIDN:AAA40738.1; PID:g202919
A;Experimental source: kidney
C;Comment: AT1 receptor consists of two closely related AT1 isoforms of angiotensin II r
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; GTP binding; receptor; transmembra
F;27-86/Domain: transmembrane #status predicted <TM1>
F;103-123/Domain: transmembrane #status predicted <TM3>
F;147-167/Domain: transmembrane #status predicted <TM4>
F;195-218/Domain: transmembrane #status predicted <TM5>
F;240-263/Domain: transmembrane #status predicted <TM6>
F;278-299/Domain: transmembrane #status predicted <TM7>
F;4,176,188/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             15.8%;  Score 341;  DB 2;  Length 359;
  Best Local Similarity   28.8%;  Pred. No. 4.5e-21;
  Matches   88;  Conservative   54;  Mismatches  126;  Indels   38;  Gaps    7;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTR-SICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              :|| :| ::|||||:  |:: : :::|  :     :||  |||:||  | :|||    |
Db         30 MIPTLYSIIFVVGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLTLPLWAVYTA 89

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTC 233
                | ||  ||:  :     |:| |: || |:||:||||||||   |       | |||
Db         90 MEYRWPFGNHLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKSRLRRTMLVAKVTC 149

Qy        234 GLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIP 293
              ::|    |  ||  | :  |::   :|| |   : :  |: |  |   ::    ||| |
Db        150 IIIWLMAGLASLPAVIHRNVYFIENTNITVCAFHYESRNSTLPIGL--GLTKNILGFLFP 207

Qy        294 FVLIIYCYAAIIRTL-NAYD-----HRWLWYVKASLLILVIFTICFAPSNIIL------- 340
              |::|:  |  | : |  ||:      :  |:: |   :  | |
Db        208 FLIILTSYTLIWKALKKAYEIQKNKPRNDDIFRIIMAIVLFFFFSWVPHQIFTFLDVLIQ 267

Qy        341 --IIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLY---------------FLMSK 382
                :||        :| :     | :|:   |:||:|  |                ::  |
Db        268 LGVIHDCKI----SDIVDTAMPITICIAYFNNCLNPLFYGFLGKKFKKYFLQLLKYIPPK 323

Qy        383 TRNHST 388
                ::||:
Db        324 AKSHSS 329


RESULT 39
JC2434
kappa opioid receptor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 19-May-2000
C;Accession: JC2434
R;Nishi, M.; Takeshima, H.; Mori, M.; Nakagawara, K.; Takeuchi, T.
Biochem. Biophys. Res. Commun. 205, 1353-1357, 1994
A;Title: Structure and chromosomal mapping of genes for the mouse kappa-opioid receptor
A;Reference number: JC2434; MUID:95100967; PMID:7802669
A;Accession: JC2434
A;Molecule type: mRNA
A;Residues: 1-380 <NIS>
A;Cross-references: DDBJ:D31663
C;Genetics:
A;Map position: 1A2-3
A;Introns: 86/2; 204/1
C;Superfamily: vertebrate rhodopsin
C;Keywords: receptor

  Query Match             15.8%;  Score 341;  DB 2;  Length 380;
  Best Local Similarity   26.6%;  Pred. No. 4.7e-21;
  Matches   93;  Conservative   62;  Mismatches  156;  Indels   38;  Gaps    8;

Qy         55 PIKTFRGAP----------PNSFEEPPFSALEGWTGATITVKIKCPEESASHLHVKNAT 103
              ||: ||| |           |||   ||      |  :     :   ::     |:  |
Db          4 PIQIFRGDPGPTCSPSACLLPNSSSWFP-----NWAESDSNGSVGSEDQQLESAHISPAI 58

Qy        104 MGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTR-SICTTVFYTNLAIADFLFC 162
                          :| |:| :|||||:  |:: ::::   |:    | ::  |||:|| |
Db         59 ---------PVIITAVYSVVFVVGLVGNSLVMFVIIRYTKMKTATNIYIFNLALADALVT 109

Qy        163 VTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRG 222
              |:||: | :|  |:| ||:|||:    | | ||: ||  |  :|::||:|: ||
Db        110 TTMPFQSAVYLM-NSWPFGDVLCKIVISIDYYNMFTSIFTLTMMSVDRYIAVCHPVKALD 168

Qy        223 LPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYF 282
                  | :   :|       :   :|         |:  |  :   :  | : |:
Db        169 FRTPLKAKIINICIWLLASSVGISAIVLGGTKVREDVDVIECL-LQFPDDEYSWWDLFMK 227

Qy        283 ISLAFFGFLIPFVLIIYCYAAIIRTLNAY---------DHRWLWYVKASLLILVIFTICF 333
              | :  | |:|| ::|| ||  :|  | :         |       |  |::: :| ||:
Db        228 ICVVVFAFVIPVLIIIVCYTLMILRLKSVRLLSGSREKDRNLRRITKLVLVVVAVFIICW 287

Qy        334 APSNIILIIHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
              | :| :::       ::|  |   |   : ||  || |:| ||  : :
Db        288 TPIHIFILVEALGSTSHSTAAL-SSYYFCIALGYTNSSLNPVLYAFLDE 335


RESULT 40
S43252
platelet-activating factor receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C;Accession: S43252
R;Bito, H.; Honda, Z.; Nakamura, M.; Shimizu, T.
```

```
Eur. J. Biochem. 221, 211-218, 1994
A;Title: Cloning, expression and tissue distribution of rat platelet-activating-factor-r
A;Reference number: S43252; MUID:94222063; PMID:8168510
A;Accession: S43252
A;Molecule type: mRNA
A;Residues: 1-341 <BIT>
A;Cross-references: GB:U04740; NID:g470384; PIDN:AAA18422.1; PID:g470385
C;Superfamily: ATP receptor P2u

  Query Match             15.8%;  Score 340;  DB 2;  Length 341;
  Best Local Similarity   30.1%;  Pred. No. 5.1e-21;
  Matches   97;  Conservative   47;  Mismatches  132;  Indels   46;  Gaps   12;

Qy        115 LIPAIYLLVFVVGVPANAVTLWM---LFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAY 171
              | | :| ::||:|| ||   ||:   |:   :     :|  || :|| || :|||  | |
Db         17 LFPIVYSVIFVLGVVANGYVLWVFATLYPSKKLNEIKIFMVNLTVADLLFLMTLPLWIVY 76

Qy        172 HLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF-------TYRGLP 224
              : |  :|:  : ||     :|: | |||:  |  |: ||| |: :|          ||:
Db         77 YSNEGDWIVHKFLCNLAGCLFFINTYCSVAFLGVITYNRYQAVAYPIKTAQATTRKRGI- 135

Qy        225 KHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPD----ITTCHDVHNTCESSSPFQLY 280
              | :||    :| ::      |       : : |    || | : |    |     ::
Db        136 --TLSLV----IWISIAATASYFLATDSTNVVPKKDGSGNITRCFE-HYEPYSVPILVVH 188

Qy        281 YFISLAFFGFLIPFVLIIYCYAAIIRTL----------NAYDHRWLWYVKASLLILVIFT 330
              ||:  ||   : | || ||   || ||               | || |     :| :|
Db        189 IFITSCFF---LVFFLIFYCNMVIIHTLLTRPVRQQRKPEVKRRALWMV---CTVLAVFV 242

Qy        331 ICFAPSNIILI---IHHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLY-FLMSKTRNH 386
              ||| | ::: :   :    | |    :   : | ||| | |  ||| :| ||  | | |
Db        243 ICFVPHHVVQLPWTLAELGYQTNFHQAINDAHQITLCLLSTNCVLDPVIYCFLTKKFRKH 302

Qy        387 STAYLTKXXNDLREQGQPSQRT 408
                  |::    :|   : |: |
Db        303 ----LSEKFYSMRSSRKCSRAT 320


Search completed: June 24, 2003, 12:06:37
Job time : 29.0319 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:05:46 ; Search time 32.5399 Seconds
(without alignments)
1356.747 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIKM.....AYLTKXXNDLREQGQPSQRT 408

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	523	24.2	425	10	US-09-782-980-80	Sequence 80, Appl
2	523	24.2	425	10	US-09-884-430-4	Sequence 4, Appli
3	471.5	21.9	359	10	US-09-943-718-6	Sequence 6, Appli
4	460	21.3	359	9	US-10-190-469-1	Sequence 1, Appli
5	460	21.3	359	9	US-10-094-417-10	Sequence 10, Appl
6	460	21.3	359	9	US-09-782-974C-76	Sequence 76, Appl
7	460	21.3	359	10	US-09-739-151-2	Sequence 2, Appli
8	428	19.8	537	9	US-10-311-956-4	Sequence 4, Appli
9	426	19.7	362	9	US-10-092-135-4	Sequence 4, Appli
10	424	19.7	362	9	US-10-092-135-3	Sequence 3, Appli
11	424	19.7	362	9	US-09-779-679-28	Sequence 28, Appl
12	413.5	19.2	339	9	US-10-251-385-182	Sequence 182, App
13	410.5	19.0	339	9	US-09-828-478-4	Sequence 4, Appli
14	410.5	19.0	339	9	US-10-251-385-32	Sequence 32, Appl
15	410.5	19.0	339	10	US-09-848-889-12	Sequence 12, Appl
16	410.5	19.0	339	10	US-09-788-133-2	Sequence 2, Appli
17	410.5	19.0	361	9	US-10-222-024-2	Sequence 2, Appli
18	410.5	19.0	361	9	US-10-251-385-78	Sequence 78, Appl
19	410.5	19.0	361	9	US-10-251-385-206	Sequence 206, App
20	410.5	19.0	367	9	US-09-828-478-6	Sequence 6, Appli
21	409.5	19.0	373	9	US-10-092-135-7	Sequence 7, Appli
22	405.5	18.8	348	10	US-09-827-937A-17	Sequence 17, Appl
23	399	18.5	373	9	US-10-092-135-6	Sequence 6, Appli
24	398	18.5	373	9	US-10-092-135-5	Sequence 5, Appli
25	397	18.4	302	9	US-10-024-494-30	Sequence 30, Appl
26	396.5	18.4	299	9	US-10-270-144-4	Sequence 4, Appli
27	395	18.3	366	9	US-09-779-679-25	Sequence 25, Appl
28	384.5	17.8	360	10	US-09-943-718-4	Sequence 4, Appli
29	382	17.7	365	9	US-09-077-173A-2	Sequence 2, Appli
30	376.5	17.5	346	9	US-09-828-478-2	Sequence 2, Appli
31	376.5	17.5	346	9	US-09-779-679-2	Sequence 2, Appli
32	376.5	17.5	346	9	US-09-779-679-26	Sequence 26, Appl
33	376.5	17.5	346	10	US-09-826-791-6	Sequence 6, Appli
34	376.5	17.5	346	10	US-09-866-230-7	Sequence 7, Appli
35	376.5	17.5	346	10	US-09-866-230-9	Sequence 9, Appli
36	375	17.4	344	9	US-10-024-494-8	Sequence 8, Appli
37	374.5	17.4	330	10	US-09-826-791-2	Sequence 2, Appli
38	372.5	17.3	337	9	US-09-828-478-5	Sequence 5, Appli
39	372.5	17.3	337	10	US-09-866-230-8	Sequence 8, Appli
40	368.5	17.1	358	10	US-09-826-508-10	Sequence 10, Appl
41	365	16.9	337	9	US-09-779-679-27	Sequence 27, Appl
42	364.5	16.9	337	9	US-10-167-192-3	Sequence 3, Appli
43	364.5	16.9	374	9	US-10-242-499-3	Sequence 3, Appli
44	361	16.7	341	9	US-10-270-587-3	Sequence 3, Appli
45	361	16.7	358	9	US-09-974-298-86	Sequence 86, Appl

ALIGNMENTS

RESULT 1
US-09-782-980-80
; Sequence 80, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, AND
; TITLE OF INVENTION: STMST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 60/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/014,348
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/086,892
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/296,208
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 09/063,950
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 09/561,381
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,810
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/087,121

```
;  PRIOR FILING DATE: 1998-05-29
;  PRIOR APPLICATION NUMBER: 09/672,721
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: 09/049,799
;  PRIOR FILING DATE: 1998-03-27
;  NUMBER OF SEQ ID NOS: 176
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
;   LENGTH: 425
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-782-980-80

  Query Match            24.2%;  Score 523;  DB 10;  Length 425;
  Best Local Similarity  32.9%;  Pred. No. 5.8e-35;
  Matches  127;  Conservative   76;  Mismatches  153;  Indels   30;  Gaps   11;

Qy         14 GTQVIKMKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPN-SFEEFPF 72
              | : : : |  |:  | ||   |  :      ||     ||  ::|    ||  :| |
Db          2 GPRRLLLVAACFSLCGPLLSARTRARRPESKATN----ATLDPRSFLLRNPNDKYEPFWE 57

Qy         73 SALEGWTGAT----ITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGV 128
                 :  :| |    :::    | :      :     ||||||  |  :|::|  |||| :
Db         58 DEEKNESGLTEYRLVSINKSSPLQKQLPAFISEDASGYLTSSWLTLFVPSVYTGVFVVSL 117

Qy        129 PAN--AVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCR 186
              | |  |: :::|  : :      |:  :|| || ||   |||||:|: :|::| ||  |||
Db        118 PLNIMAIVVFILKMKVKK-PAVVYMLHLATADVLFVSVLPFKISYYFSGSDWQFGSELCR 176

Qy        187 ATTVIFYGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
                |  || ||| |||:  |||:|:||:|:|    ::| | : ::   ||  :||
Db        177 FVTAAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASF---TCLAIWALAIAG 233

Qy        244 MLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAA 303
              ::|  : :|   :  :|||||| |       :  ||| : :   | :| ::   || :
Db        234 VVPLVLKEQTIQVPGLNITTCHDVLNETLLEG-YYAYYFSAFSAVFFFVPLIISTVCYVS 292

Qy        304 IIRTLN-------AYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANY-YYNNTDGL 355
              ||| |:       :   | |:    |  :  || ||| |:|::|| |::   : : |:
Db        293 IIRCLSSSAVANRSKKSRALF---LSAAVFCIFIICFGPTNVLLIAHYSFLSHTSTTEAA 349

Qy        356 YFIYLIALCLGSLNSCLDPFLYFLMS 381
              || ||: :|: |::||:|| :|:  |
Db        350 YFAYLLCVCVSSISSCIDPLIYYYAS 375


RESULT 2
US-09-884-430-4
; Sequence 4, Application US/09884430
; Patent No. US20020151046A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria Alexandra
;  APPLICANT:  Santiago, Immaculada Silos
;  TITLE OF INVENTION: 52871, A NOVEL HUMAN G PROTEIN COUPLED
;  TITLE OF INVENTION:  RECEPTOR AND USES THEREOF
;  FILE REFERENCE: MNI-165
;  CURRENT APPLICATION NUMBER: US/09/884,430
;  CURRENT FILING DATE:  2001-06-18
;  PRIOR APPLICATION NUMBER: USSN 60/212,331
;  PRIOR FILING DATE: 2000-06-16
;  PRIOR APPLICATION NUMBER: USSN 60/269,758
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 425
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-884-430-4

  Query Match            24.2%;  Score 523;  DB 10;  Length 425;
  Best Local Similarity  32.9%;  Pred. No. 5.8e-35;
  Matches  127;  Conservative   76;  Mismatches  153;  Indels   30;  Gaps   11;

Qy         14 GTQVIKMKALIFAAAGLLLLLPTFCQSGMENDTNNLAKPTLPIKTFRGAPPN-SFEEFPF 72
              | : : : |  |:  | ||   |  :      ||     ||  ::|    ||  :| |
Db          2 GPRRLLLVAACFSLCGPLLSARTRARRPESKATN----ATLDPRSFLLRNPNDKYEPFWE 57

Qy         73 SALEGWTGAT----ITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGV 128
                 :  :| |    :::    | :      :     ||||||  |  :|::|  |||| :
Db         58 DEEKNESGLTEYRLVSINKSSPLQKQLPAFISEDASGYLTSSWLTLFVPSVYTGVFVVSL 117

Qy        129 PAN--AVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCR 186
              | |  |: :::|  : :      |:  :|| || ||   |||||:|: :|::| ||  |||
Db        118 PLNIMAIVVFILKMKVKK-PAVVYMLHLATADVLFVSVLPFKISYYFSGSDWQFGSELCR 176

Qy        187 ATTVIFYGNMYCSILLLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
                |  || ||| |||:  |||:|:||:|:|    ::| | : ::   ||  :||
Db        177 FVTAAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASF---TCLAIWALAIAG 233

Qy        244 MLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAA 303
              ::|  : :|   :  :|||||| |       :  ||| : :   | :| ::   || :
Db        234 VVPLVLKEQTIQVPGLNITTCHDVLNETLLEG-YYAYYFSAFSAVFFFVPLIISTVCYVS 292

Qy        304 IIRTLN-------AYDHRWLWYVKASLLILVIFTICFAPSNIILIIHHANY-YYNNTDGL 355
              ||| |:       :   | |:    |  :  || ||| |:|::|| |::   : : |:
Db        293 IIRCLSSSAVANRSKKSRALF---LSAAVFCIFIICFGPTNVLLIAHYSFLSHTSTTEAA 349

Qy        356 YFIYLIALCLGSLNSCLDPFLYFLMS 381
              || ||: :|: |::||:|| :|:  |
Db        350 YFAYLLCVCVSSISSCIDPLIYYYAS 375


RESULT 3
US-09-943-718-6
; Sequence 6, Application US/09943718
; Patent No. US20020103361A1
;  GENERAL INFORMATION:
;       APPLICANT: Huffine, Constance F.
;                  Rossi, Devora L.
;                  Capone, Myriam
;                  Hedrick, Joseph A.
;                  Vicari, Alain
;                  Gorman, Daniel M.
;                  Zlotnik, Albert
;       TITLE OF INVENTION: Mammalian Chemokines; Receptors;
;                           Reagents; Uses
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DNAX Research Institute
;            STREET: 901 California Avenue
;            CITY: Palo Alto
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94304-1104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/943,718
;            FILING DATE: 30-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/009,817
;            FILING DATE: 20-APR-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ching, Edwin P.
;            REGISTRATION NUMBER: 34,090
;            REFERENCE/DOCKET NUMBER: DX0588K
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (650)852-9196
;            TELEFAX: (650)496-1200
;  INFORMATION FOR SEQ ID NO: 6:
;       SEQUENCE CHARACTERISTICS:
```

```
;             LENGTH: 359 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-943-718-6

  Query Match             21.9%;  Score 471.5;  DB 10;  Length 359;
  Best Local Similarity   36.6%;  Pred. No. 7.7e-31;
  Matches  102;  Conservative   47;  Mismatches  111;  Indels   19;  Gaps    4;

Qy        112 STKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSICTTVFYTNLAIADFLFCVTLPFKIAY 171
              | :|:||:| ||  ||:||| : ||:|  |   : :|:   |||:|| |  : || ::||
Db         51 SPQLVPALYGLVVAVGLPANGLALWVLATRVPRLPSTILLMNLAVADLLLALVLPPRLAY 110

Qy        172 HLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALV 231
              || |  | |||  ||  |   ||:|| |:|||| :|::||||:|||   | |
Db        111 HLRGQRWPFGEAACRVATAALYGHMYGSVLLLAAVSLDRYLALVHPLRARALRGQRLTTG 170

Qy        232 TCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHNTC---------ESSSPFQLYYF 282
              | : | :     ||  : :| : |: | |  |  |   |         |  |   | :
Db        171 LCLVAWLSAATLALPLTLHRQNFRLLAP-IACC--VMMRCPWLSRTPTGERPSSAWLSWA 227

Qy        283 ISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWYVKASLLILVIFTICFAPSNIILII 342
              ||        | : :  ||    :| | |    |:   :: : |:|       | |||::|::
Db        228 ASL-------PLLAMGLCYGTTLRALAANGQRYSHALRLTALVLFSAVASFTPSNVLLVL 280

Qy        343 HHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMS 381
              |::|       ||  |: :| | :||||:|||:|: :|
Db        281 HYSNPSPEAWGNLYGAYVPSLALSTLNSCVDPFIYYYVS 319


RESULT 4
US-10-190-469-1
; Sequence 1, Application US/10190469
; Publication No. US20030013155A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  TITLE OF INVENTION: 14400 Receptor, A No. US20030013155A1el G-Protein Coupled Recepto
;  FILE REFERENCE: 5800-7, 035800/169196
;  CURRENT APPLICATION NUMBER: US/10/190,469
;  CURRENT FILING DATE:  2002-07-05
;  PRIOR APPLICATION NUMBER: US/09/137,063A
;  PRIOR FILING DATE: 1999-08-20
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-190-469-1

  Query Match             21.3%;  Score 460;  DB 9;  Length 359;
  Best Local Similarity   34.4%;  Pred. No. 6.7e-30;
  Matches  100;  Conservative   54;  Mismatches  125;  Indels   12;  Gaps    7;

Qy        101 NATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFR--TRSICTTVFYTNLAIAD 158
              |||:  | :       :| :| ||  | :| |  :||:|  |    ||  : :|  ||:: |
Db         11 NATLQMLRNPAIAVALPVVYSLVAAVSIPGNLFSLWVLCRRMGPRS-PSVIFMINLSVTD 69

Qy        159 FLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF 218
               :   |||:| || | ::|||| :||   || || ||| ||| : |||: |:| :::|
Db         70 LMLASVLPFQIYYHCNRHHWVFGVLLCNVVTVAFYANMYSSILTMTCISVERFLGVLYPL 129

Qy        219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHN-TCESSSPF 277
              : :   :  ||:  |   |  :   : |        | :    | || ||   |   |
Db        130 SSKRWRRRRYAVAACAGTWLLLLTALSPLARTDLTYPVHALGIITCFDVLKWTMLPSVAM 189

Qy        278 QLYYFISLAFFGFLIPFVLIIYCYAA----IIRTLNAYD-HRWLWYVKASLLILVIFTIC 332
              :  ::    ||||||: : || |   .::||  |:   :    |  : ::|: |  |
Db        190 WAVFLFTIFILLFLIPFVITVACYTATILKLLRTEEAHGREQRRRAVGLAAVVLLAFVTC 249

Qy        333 FAPSNIILIIH-HANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
              |||:| :|: |  :  :|  :   | :| : |||  ||:|||||:|:  |:
Db        250 FAPNNFVLLAHIVSRLFYGKS--YYHVYKLTLCLSCLNNCLDPFVYYFASR 298


RESULT 5
US-10-094-417-10
; Sequence 10, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
;  APPLICANT: Tian, Hui
;  APPLICANT:  Zhao, Jiagang
;  APPLICANT:  Chen, Jin-Long
;  APPLICANT:  Cutler, Gene
;  APPLICANT:  Tularik Inc.
;  TITLE OF INVENTION: No. US20030045685A1el Receptors
;  FILE REFERENCE: 018781-008110US
;  CURRENT APPLICATION NUMBER: US/10/094,417
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: US 09/802,803
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: US 60/276,649
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR341
US-10-094-417-10

  Query Match             21.3%;  Score 460;  DB 9;  Length 359;
  Best Local Similarity   34.4%;  Pred. No. 6.7e-30;
  Matches  100;  Conservative   54;  Mismatches  125;  Indels   12;  Gaps    7;

Qy        101 NATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFR--TRSICTTVFYTNLAIAD 158
              |||:  | :       :| :| ||  | :| |  :||:|  |    ||  : :|  ||:: |
Db         11 NATLQMLRNPAIAVALPVVYSLVAAVSIPGNLFSLWVLCRRMGPRS-PSVIFMINLSVTD 69

Qy        159 FLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF 218
               :   |||:| || | ::|||| :||   || || ||| ||| : |||: |:| :::|
Db         70 LMLASVLPFQIYYHCNRHHWVFGVLLCNVVTVAFYANMYSSILTMTCISVERFLGVLYPL 129

Qy        219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHN-TCESSSPF 277
              : :   :  ||:  |   |  :   : |        | :    | || ||   |   |
Db        130 SSKRWRRRRYAVAACAGTWLLLLTALSPLARTDLTYPVHALGIITCFDVLKWTMLPSVAM 189

Qy        278 QLYYFISLAFFGFLIPFVLIIYCYAA----IIRTLNAYD-HRWLWYVKASLLILVIFTIC 332
              :  ::    ||||||: : || |   .::||  |:   :    |  : ::|: |  |
Db        190 WAVFLFTIFILLFLIPFVITVACYTATILKLLRTEEAHGREQRRRAVGLAAVVLLAFVTC 249

Qy        333 FAPSNIILIIH-HANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
              |||:| :|: |  :  :|  :   | :| : |||  ||:|||||:|:  |:
Db        250 FAPNNFVLLAHIVSRLFYGKS--YYHVYKLTLCLSCLNNCLDPFVYYFASR 298


RESULT 6
US-09-782-974C-76
; Sequence 76, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
;  APPLICANT: Vogeli, Gabriel
;  APPLICANT:  Lind, Peter
;  APPLICANT:  Wood, Linda S.
;  APPLICANT:  Parodi, Luis A.
;  TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
;  FILE REFERENCE: 411USPHRM311
;  CURRENT APPLICATION NUMBER: US/09/782,974C
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: 60/165,838
;  PRIOR FILING DATE: 1999-11-16
;  PRIOR APPLICATION NUMBER: 09/714,449
```

```
;  PRIOR FILING DATE: 2000-11-16
;  PRIOR APPLICATION NUMBER: 60/198,568
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: 60/166,071
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/166,678
;  PRIOR FILING DATE: 1999-11-19
;  PRIOR APPLICATION NUMBER: 60/173,396
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/184,129
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/185,421
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/185,554
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/186,530
;  PRIOR FILING DATE: 2000-03-02
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 192
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-782-974C-76

  Query Match             21.3%;  Score 460;  DB 9;  Length 359;
  Best Local Similarity   34.4%;  Pred. No. 6.7e-30;
  Matches  100;  Conservative   54;  Mismatches  125;  Indels   12;  Gaps    7;

Qy        101 NATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFR--TRSICTTVFYTNLAIAD 158
              |||:  | :      :| :| ||  | :| |  :||:|  |   ||  : :|  ||:: |
Db         11 NATLQMLRNPAIAVALPVVYSLVAAVSIPGNLFSLWVLCRRMGPRS-PSVIFMINLSVTD 69

Qy        159 FLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF 218
               :    |||:| || | ::|||| :||   || || ||| ||| : |||: |:| :::|
Db         70 LMLASVLPFQIYYHCNRHHWVFGVLLCNVVTVAFYANMYSSILTMTCISVERFLGVLYPL 129

Qy        219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHN-TCESSSPF 277
              : :   :  ||:  |   |  :   : |       | :    | || ||   |   |
Db        130 SSKRWRRRRYAVAACAGTWLLLLTALSPLARTDLTYPVHALGIITCFDVLKWTMLPSVAM 189

Qy        278 QLYYFISLAFFGFLIPFVLIIYCYAA----IIRTLNAYD-HRWLWYVKASLLILVIFTIC 332
                :  ::    ||||||: : || |    ::||  |:   :    |  : ::|: |  |
Db        190 WAVFLFTIFILLFLIPFVITVACYTATILKLLRTEEAHGREQRRRAVGLAAVVLLAFVTC 249

Qy        333 FAPSNIILIIH-HANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
              |||:| :|: |  :  :|  :   | :| : |||  ||:|||||:|:  |:
Db        250 FAPNNFVLLAHIVSRLFYGKS--YYHVYKLTLCLSCLNNCLDPFVYYFASR 298


RESULT 7
US-09-739-151-2
; Sequence 2, Application US/09739151
; Patent No. US20010029032A1
; GENERAL INFORMATION:
;  APPLICANT:  Yuan Zhu
;  APPLICANT:  Xiaotong Li
;  APPLICANT:  Lisa Vawter
;  TITLE OF INVENTION:  PAUL, A G-PROTEIN COUPLED RECEPTOR
;  TITLE OF INVENTION:  POLYPEPTIDE
;  FILE REFERENCE:  GP-70567-C1
;  CURRENT APPLICATION NUMBER: US/09/739,151
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER:  09/413,534
;  PRIOR FILING DATE:  1999-10-06
;  PRIOR APPLICATION NUMBER:  60/103,789
;  PRIOR FILING DATE:  1998-10-09
;  NUMBER OF SEQ ID NOS:  4
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-739-151-2

  Query Match             21.3%;  Score 460;  DB 10;  Length 359;
  Best Local Similarity   34.4%;  Pred. No. 6.7e-30;
  Matches  100;  Conservative   54;  Mismatches  125;  Indels   12;  Gaps    7;

Qy        101 NATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFR--TRSICTTVFYTNLAIAD 158
              |||:  | :      :| :| ||  | :| |  :||:|  |   ||  : :|  ||:: |
Db         11 NATLQMLRNPAIAVALPVVYSLVAAVSIPGNLFSLWVLCRRMGPRS-PSVIFMINLSVTD 69

Qy        159 FLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPF 218
               :    |||:| || | ::|||| :||   || || ||| ||| : |||: |:| :::|
Db         70 LMLASVLPFQIYYHCNRHHWVFGVLLCNVVTVAFYANMYSSILTMTCISVERFLGVLYPL 129

Qy        219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDITTCHDVHN-TCESSSPF 277
              : :   :  ||:  |   |  :   : |       | :    | || ||   |   |
Db        130 SSKRWRRRRYAVAACAGTWLLLLTALSPLARTDLTYPVHALGIITCFDVLKWTMLPSVAM 189

Qy        278 QLYYFISLAFFGFLIPFVLIIYCYAA----IIRTLNAYD-HRWLWYVKASLLILVIFTIC 332
                :  ::    ||||||: : || |    ::||  |:   :    |  : ::|: |  |
Db        190 WAVFLFTIFILLFLIPFVITVACYTATILKLLRTEEAHGREQRRRAVGLAAVVLLAFVTC 249

Qy        333 FAPSNIILIIH-HANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
              |||:| :|: |  :  :|  :   | :| : |||  ||:|||||:|:  |:
Db        250 FAPNNFVLLAHIVSRLFYGKS--YYHVYKLTLCLSCLNNCLDPFVYYFASR 298


RESULT 8
US-10-311-956-4
; Sequence 4, Application US/10311956
; Publication No. US20030109482A1
; GENERAL INFORMATION:
;  APPLICANT: Bayer AG
;  APPLICANT:  Bull, Christof
;  TITLE OF INVENTION: REGULATION OF HUMAN P2Y8-LIKE G PROTEIN-COUPLED RECEPTOR
;  FILE REFERENCE: Lio095 foreign countries
;  CURRENT APPLICATION NUMBER: US/10/311,956
;  CURRENT FILING DATE:  2002-12-20
;  PRIOR APPLICATION NUMBER: 60/214,257
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: 60/254,878
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 537
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-311-956-4

  Query Match             19.8%;  Score 428;  DB 9;  Length 537;
  Best Local Similarity   36.0%;  Pred. No. 4.2e-27;
  Matches  105;  Conservative   41;  Mismatches  102;  Indels   44;  Gaps   12;

Qy        115 LIPAIYLLVFVVGVPANAVTLWMLFFRTRS-ICTTVFYTNLAIADFLFCVTLPFKIAYHL 173
              |:|  |  ||:||:| |   :|:   : |    |||:  |||::| |: ::||  : |:
Db         45 LLPVSYSAVFMVGLPLNIAAMWIFIAKMRPWNPTTVYMFNLALSDTLYVLSLPTLVYYYA 104

Qy        174 NGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFT--YRGLPKHTYALV 231
              : ||| ||||||:    :|| |:| ||| | |||::||  : || |   |   || |  |
Db        105 DKNNWPFGEVLCKLVRFLFYANLYSSILFLTCISVHRYRGVCHPITSLRRMNAKHAY--V 162

Qy        232 TCGLVWATVFLYMLPFFILKQEYYLVQPDI--TTCHDVHNTCESSSPFQLYYFISLAFFG 289
              | ||| :| | ::|  |    :  | | :  | |||           |  |   | |
Db        163 ICALVWLSVTLCLVPNLI----FVTVSPKVKNTICHDT----TRPEDFARYVEYSTAIMC 214

Qy        290 FL--IPFVLIIYCYAAIIR-----------TLNAYDHRWLWYVKASLLILVIFTICFAP 335
              |  || ::|  ||  : |           || :|  |    :|  : ::: | ||| |
Db        215 LLFGIPCLIIAGCYGLMTRELMKPIVSGNQQTLPSYKKR---SIKTIIFVMIAFAICFMP 271

Qy        336 SNIILIIHHANYYYNNTDGL--YFI------YLIALCLGSLNSCLDPFLYFL 379
              :|   :    |||    |:  | :      | :   | | |||:|| ||||
```

```
Db        272 FHITRTL----YYYARLLGIKCYALNVINVTYKVTRPLASANSCIDPILYFL 319


RESULT 9
US-10-092-135-4
; Sequence 4, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
;  TITLE OF INVENTION:  HGPRBMY27
;  FILE REFERENCE: D0134.NP
;  CURRENT APPLICATION NUMBER: US/10/092,135
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: US 60/273,808
;  PRIOR FILING DATE: 2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/278,983
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 362
;   TYPE: PRT
;   ORGANISM: MELEAGRIS GALLOPAVO
US-10-092-135-4

  Query Match           19.7%;  Score 426;  DB 9;  Length 362;
  Best Local Similarity   30.4%;  Pred. No. 4e-27;
  Matches  101;  Conservative   59;  Mismatches  132;  Indels   40;  Gaps    9;

Qy        105 GYLTSSLSTK-----------LIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYT 152
              |:    : |||           :| :|:|||: |   |:| :|| |  |     :|:
Db         21 GWAAGNASTKCSLTKTGFQFYYLPTVYILVFITGFLGNSVAIWMFVFHMRPWSGISVYMF 80

Qy        153 NLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYL 212
              |||:||||: :|||  | |: |  :|:||:|:|:    ||: |:| ||| | |||::||
Db         81 NLALADFLYVLTLPALIFYYFNKTDWIFGDVMCKLQRFIFHVNLYGSILFLTCISVHRYT 140

Qy        213 AIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY----LVQPDITTCHDVH 268
               :|||    | |   |:    |||| |   : |       :|    : :    ||:|
Db        141 GVVHPLKSLGRLKKKNAVYVSSLVWALVVAVIAPIL-----FYSGTGVRRNKTITCYD-- 193

Qy        269 NTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKA 321
              | :       | :    | | |||::|: ||  |:: |   |         : :: |
Db        194 TTADEYLRSYFVYSMCTTVFMFCIPFIVILGCYGLIVKALIYKDLDNSPLRRKSIYLV-- 251

Qy        322 SLLILVIFTICFAPSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDP 374
               :::| :| : : | :::  ::  |   :          |:|  | :   | |||||:||
Db        252 -IIVLTVFAVSYLPFHVMKTLNLRARLDFQTPQMCAFNDKVYATYQVTRGLASLNSCVDP 310

Qy        375 FLYFLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
               ||||   |     :  |:  :   |   |:
Db        311 ILYFLAGDTFRRRLSRATRKSSRRSEPNVQSK 342


RESULT 10
US-10-092-135-3
; Sequence 3, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
;  TITLE OF INVENTION:  HGPRBMY27
;  FILE REFERENCE: D0134.NP
;  CURRENT APPLICATION NUMBER: US/10/092,135
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: US 60/273,808
;  PRIOR FILING DATE: 2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/278,983
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 362
;   TYPE: PRT
;   ORGANISM: GALLUS GALLUS
US-10-092-135-3

  Query Match           19.7%;  Score 424;  DB 9;  Length 362;
  Best Local Similarity   28.9%;  Pred. No. 5.8e-27;
  Matches  101;  Conservative   59;  Mismatches  143;  Indels   46;  Gaps    9;

Qy         77 GWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLW 136
              ||     | |    :    :                :| :|:|||: |   |:| :|
Db         21 GWAAGNATTKCSLTKTGFQFYY----------------LPTVYILVFITGFLGNSVAIW 63

Qy        137 MLFFRTRSIC-TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGN 195
              |  |  |     :|:  |||:||||: :|||  | |: |  :|:||:|:|:    ||: |
Db         64 MFVFHMRPWSGISVYMFNLALADFLYVLTLPALIFYYFNKTDWIFGDVMCKLQRFIFHVN 123

Qy        196 MYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY 255
              :| ||| | |||::||  :|||    | |   |:    |||| |   : |       :|
Db        124 LYGSILFLTCISVHRYTGVVHPLKSLGRLKKKNAVYVSSLVWALVVAVIAPIL-----FY 178

Qy        256 ----LVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAY 311
                  : :   ||:|   | :    | :   | | |||::|: ||  |:: |
Db        179 SGTGVRRNKTITCYD--TTADEYLRSYFVYSMCTTVFMFCIPFIVILGCYGLIVKALIYK 236

Qy        312 D-------HRWLWYVKASLLILVIFTICFAPSNIILIIH-HANYYYNN------TDGLYF 357
              |       : :: |   :::| :| : : | :::  ::  |   :         | :|
Db        237 DLDNSPLRRKSIYLV---IIVLTVFAVSYLPFHVMKTLNLRARLDFQTPQMCAFNDKVYA 293

Qy        358 IYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
              | :   | |||||:|| ||||   |     :  |:  :   |    |:
Db        294 TYQVTRGLASLNSCVDPILYFLAGDTFRRRLSRATRKSSRRSEPNVQSK 342


RESULT 11
US-09-779-679-28
; Sequence 28, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond J
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Mishra, Vishnu S
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Ballinger, Robert
;  APPLICANT:  Vernet, Corine A
;  APPLICANT:  Li, Li
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Andrew, David P
;  APPLICANT:  Mezes, Peter S
;  TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sa
;  FILE REFERENCE: 15966-661
;  CURRENT APPLICATION NUMBER: US/09/779,679
;  CURRENT FILING DATE:  2001-02-08
;  PRIOR APPLICATION NUMBER: USSN 60/181045
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: USSN 60/183191
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: USSN 60/180,929
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: USSN 60/219758
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: USSN 60/181339
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: USSN 60/181344
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: USSN 60/221341
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: USSN 60/181392
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: USSN 60/219585
;  PRIOR FILING DATE: 2000-07-20
```

```
;  PRIOR APPLICATION NUMBER: USSN 60/181157
;  PRIOR FILING DATE: 2000-02-09
;  NUMBER OF SEQ ID NOS: 65
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 362
;   TYPE: PRT
;   ORGANISM: Gallus gallus
US-09-779-679-28

  Query Match            19.7%;  Score 424;  DB 9;  Length 362;
  Best Local Similarity  28.9%;  Pred. No. 5.8e-27;
  Matches  101;  Conservative   59;  Mismatches  143;  Indels   46;  Gaps    9;

Qy         77 GWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYLLVFVVGVPANAVTLW 136
              ||     | |    :     :                 :| :|:|||: |   |:| :|
Db         21 GWAAGNATTKCSLTKTGFQFYY----------------LPTVYILVFITGFLGNSVAIW 63

Qy        137 MLFFRTRSIC-TTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGN 195
              |  |  |     :|:  |||:||||: :|||  | |: |  :|:||:|:|:    ||: |
Db         64 MFVFHMRPWSGISVYMFNLALADFLYVLTLPALIFYYFNKTDWIFGDVMCKLQRFIFHVN 123

Qy        196 MYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY 255
              :| ||| | |||::||  :|||    |  |   |:    |||| |   : |        :|
Db        124 LYGSILFLTCISVHRYTGVVHPLKSLGRLKKKNAVYVSSLVWALVVAVIAPIL-----FY 178

Qy        256 ----LVQPDITTCHDVHNTCESSSPFQLYYFISLAFFGFLIPFVLIIYCYAAIIRTLNAY 311
                  : :    ||:|   | :         | :    | | |||::|: ||  |:: |
Db        179 SGTGVRRNKTITCYD--TTADEYLRSYFVYSMCTTVFMFCIPFIVILGCYGLIVKALIYK 236

Qy        312 D-------HRWLWYVKASLLILVIFTICFAPSNIILIIH-HANYYYNN------TDGLYF 357
              |        : :: |   :::| :| : : | :::  ::  |   :          | :|
Db        237 DLDNSPLRRKSIYLV---IIVLTVFAVSYLPFHVMKTLNLRARLDFQTPQMCAFNDKVYA 293

Qy        358 IYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
               | :   | |||||:|| ||||   |     :  |:  :   |     |:
Db        294 TYQVTRGLASLNSCVDPILYFLAGDTFRRRLSRATRKSSRRSEPNVQSK 342


RESULT 12
US-10-251-385-182
; Sequence 182, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
;  TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
;  TITLE OF INVENTION:  Protein-Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/10/251,385
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US/09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 182
;   LENGTH: 339
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-251-385-182

  Query Match            19.2%;  Score 413.5;  DB 9;  Length 339;
  Best Local Similarity  31.0%;  Pred. No. 3.9e-26;
  Matches  108;  Conservative   49;  Mismatches  130;  Indels   61;  Gaps   13;

Qy         62 APPNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYL 121
              |||     |  :  |           :| :|                 : |    |  : ||
Db          7 APPGLITNFSLATAE----------QCGQE---------------TPLENMLFASFYL 39

Qy        122 LVFVVGVPANAVTLWMLFFRTRSICT--TVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV 179
              | |:: :  | : || || |     |   ||  :||:||    : || :: || :||:|

Db         40 LDFILALVGNTLALW-LFIRDHKSGTPANVFLMHLAVADLSCVLVLPTRLVYHFSGNHWP 98

Qy        180 FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
              |||: || |  :|| ||| ||  | ||| :|:||||||     | :  || : |  :|
Db         99 FGEIACRLTGFLFYLNMYASIYFLTCISADRFLAIVHPVKSLKLRRPLYAHLACAFLWVV 158

Qy        240 VFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLY-------YFISLAFFGFLI 292
              | : | |  :  |         : | | |   |       |||          :|||   |
Db        159 VAVAMAPLLVSPQ-------TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203

Qy        293 PFVLIIYCYAAIIRTLN---AYDHRWLWYVKASL-LILVIFTICFAPSNI---ILIIHHA 345
              ||:  : ||  |||:|       : |   |  : ::| || :|| | ::   : ::|:
Db        204 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAKRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 263

Qy        346 NYYYNNTDG--LYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
              ::  :      |      |  || |||  ||| :|| :::   |:   |
Db        264 SHGASCATQRILALANRITSCLTSLNGALDPIMYFFVAEKFRHALCNL 311


RESULT 13
US-09-828-478-4
; Sequence 4, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
;  APPLICANT: Xiao, Yonghong
;  TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
;  TITLE OF INVENTION:  Protein
;  FILE REFERENCE: 04974.00458
;  CURRENT APPLICATION NUMBER: US/09/828,478
;  CURRENT FILING DATE:  2001-04-09
;  PRIOR APPLICATION NUMBER: 60/195,196
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: 60/254,876
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 339
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-828-478-4

  Query Match            19.0%;  Score 410.5;  DB 9;  Length 339;
  Best Local Similarity  30.7%;  Pred. No. 6.8e-26;
  Matches  107;  Conservative   48;  Mismatches  132;  Indels   61;  Gaps   12;

Qy         62 APPNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYL 121
              |||     |  :  |           :| :|                 : |    |  : ||
Db          7 APPGLITNFSLATAE----------QCGQE---------------TPLENMLFASFYL 39

Qy        122 LVFVVGVPANAVTLWMLFFRTRSICT--TVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV 179
              | |:: :  | : || || |     |   ||  :||:||    : || :: || :||:|

Db         40 LDFILALVGNTLALW-LFIRDHKSGTPANVFLMHLAVADLSCVLVLPTRLVYHFSGNHWP 98

Qy        180 FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
              |||: || |  :|| ||| ||  | ||| :|:||||||     | :  || : |  :|
Db         99 FGEIACRLTGFLFYLNMYASIYFLTCISADRFLAIVHPVKSLKLRRPLYAHLACAFLWVV 158

Qy        240 VFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLY-------YFISLAFFGFLI 292
              | : | |  :  |         : | | |   |       |||          :|||   |
Db        159 VAVAMAPLLVSPQ-------TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203

Qy        293 PFVLIIYCYAAIIRT----LNAYDHRWLWYVKASLLILVIFTICFAPSNI---ILIIHHA 345
              ||:  : ||  |||:     |          |:   ::| || :|| | ::   : ::|:
Db        204 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 263

Qy        346 NYYYNNTDG--LYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
              ::  :      |      |  || |||  ||| :|| :::   |:   |
Db        264 SHGASCATQRILALANRITSCLTSLNGALDPIMYFFVAEKFRHALCNL 311


RESULT 14
US-10-251-385-32
```

```
; Sequence 32, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
;  TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
;  TITLE OF INVENTION:  Protein-Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/10/251,385
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US/09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 339
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-251-385-32

  Query Match             19.0%;  Score 410.5;  DB 9;  Length 339;
  Best Local Similarity   30.7%;  Pred. No. 6.8e-26;
  Matches  107;  Conservative   48;  Mismatches  132;  Indels   61;  Gaps   12;

Qy         62 APPNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYL 121
              |||      |  :  |           :| :|                 : |   |  : ||
Db          7 APPGLITNFSLATAE----------QCGQE---------------TPLENMLFASFYL 39

Qy        122 LVFVVGVPANAVTLWMLFFRTRSICT--TVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV 179
              | |:: :  | : || || |     |   ||  :||:||    : || :: || :||:|
Db         40 LDFILALVGNTLALW-LFIRDHKSGTPANVFLMHLAVADLSCVLVLPTRLVYHFSGNHWP 98

Qy        180 FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
              |||: || |  :|| ||| ||  | ||| :|:||||||     | :  || : |  :|
Db         99 FGEIACRLTGFLFYLNMYASIYFLTCISADRFLAIVHPVKSLKLRRPLYAHLACAFLWVV 158

Qy        240 VFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLY-------YFISLAFFGFLI 292
              | : | |  :  |         : | | |    |       |||          :|||   |
Db        159 VAVAMAPLLVSPQ-------TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203

Qy        293 PFVLIIYCYAAIIRT----LNAYDHRWLWYVKASLLILVIFTICFAPSNI---ILIIHHA 345
              ||:  : ||  |||:    |          |:   ::| || :|| | ::   : ::|:
Db        204 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 263

Qy        346 NYYYNNTDG--LYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
              .  ::  :      |     |  || |||  ||| :|| :::   |:   |
Db        264 SHGASCATQRILALANRITSCLTSLNGALDPIMYFFVAEKFRHALCNL 311


RESULT 15
US-09-848-889-12
; Sequence 12, Application US/09848889
; Patent No. US20020025555A1
; GENERAL INFORMATION:
;  APPLICANT: Au-Young, Janice
;  APPLICANT:  Guegler, Karl J.
;  APPLICANT:  Cheng, Muzong
;  TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER
;  FILE REFERENCE: PC-0042 CIP
;  CURRENT APPLICATION NUMBER: US/09/848,889
;  CURRENT FILING DATE:  2001-05-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PERL Program
; SEQ ID NO 12
;   LENGTH: 339
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020025555A1 g992700
US-09-848-889-12

  Query Match             19.0%;  Score 410.5;  DB 10;  Length 339;
  Best Local Similarity   30.7%;  Pred. No. 6.8e-26;
  Matches  107;  Conservative   48;  Mismatches  132;  Indels   61;  Gaps   12;

Qy         62 APPNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYL 121
              |||      |  :  |           :| :|                 : |   |  : ||
Db          7 APPGLITNFSLATAE----------QCGQE---------------TPLENMLFASFYL 39

Qy        122 LVFVVGVPANAVTLWMLFFRTRSICT--TVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV 179
              | |:: :  | : || || |     |   ||  :||:||    : || :: || :||:|
Db         40 LDFILALVGNTLALW-LFIRDHKSGTPANVFLMHLAVADLSCVLVLPTRLVYHFSGNHWP 98

Qy        180 FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
              |||: || |  :|| ||| ||  | ||| :|:||||||     | :  || : |  :|
Db         99 FGEIACRLTGFLFYLNMYASIYFLTCISADRFLAIVHPVKSLKLRRPLYAHLACAFLWVV 158

Qy        240 VFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLY-------YFISLAFFGFLI 292
              | : | |  :  |         : | | |    |       |||          :|||   |
Db        159 VAVAMAPLLVSPQ-------TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203

Qy        293 PFVLIIYCYAAIIRT----LNAYDHRWLWYVKASLLILVIFTICFAPSNI---ILIIHHA 345
              ||:  : ||  |||:    |          |:   ::| || :|| | ::   : ::|:
Db        204 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 263

Qy        346 NYYYNNTDG--LYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
              .  ::  :      |     |  || |||  ||| :|| :::   |:   |
Db        264 SHGASCATQRILALANRITSCLTSLNGALDPIMYFFVAEKFRHALCNL 311


RESULT 16
US-09-788-133-2
; Sequence 2, Application US/09788133
; Patent No. US20020052001A1
; GENERAL INFORMATION:
;  APPLICANT: GLAXO GROUP LTD
;  TITLE OF INVENTION: ASSAY
;  FILE REFERENCE: P79011
;  CURRENT APPLICATION NUMBER: US/09/788,133
;  CURRENT FILING DATE:  2001-02-20
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 339
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-788-133-2

  Query Match             19.0%;  Score 410.5;  DB 10;  Length 339;
  Best Local Similarity   30.7%;  Pred. No. 6.8e-26;
  Matches  107;  Conservative   48;  Mismatches  132;  Indels   61;  Gaps   12;

Qy         62 APPNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYL 121
              |||      |  :  |           :| :|                 : |   |  : ||
Db          7 APPGLITNFSLATAE----------QCGQE---------------TPLENMLFASFYL 39

Qy        122 LVFVVGVPANAVTLWMLFFRTRSICT--TVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV 179
              | |:: :  | : || || |     |   ||  :||:||    : || :: || :||:|
Db         40 LDFILALVGNTLALW-LFIRDHKSGTPANVFLMHLAVADLSCVLVLPTRLVYHFSGNHWP 98

Qy        180 FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
              |||: || |  :|| ||| ||  | ||| :|:||||||     | :  || : |  :|
Db         99 FGEIACRLTGFLFYLNMYASIYFLTCISADRFLAIVHPVKSLKLRRPLYAHLACAFLWVV 158

Qy        240 VFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLY-------YFISLAFFGFLI 292
              | : | |  :  |         : | | |    |       |||          :|||   |
Db        159 VAVAMAPLLVSPQ-------TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203

Qy        293 PFVLIIYCYAAIIRT----LNAYDHRWLWYVKASLLILVIFTICFAPSNI---ILIIHHA 345
              ||:  : ||  |||:    |          |:   ::| || :|| | ::   : ::|:
Db        204 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 263

Qy        346 NYYYNNTDG--LYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
```

```
Db        264 SHGASCATQRILALANRITSCLTSLNGALDPIMYFFVAEKFRHALCNL 311


RESULT 17
US-10-222-024-2
; Sequence 2, Application US/10222024
; Publication No. US20030104487A1
; GENERAL INFORMATION:
;  APPLICANT: Pfizer Ltd. (GB)
;  APPLICANT:  Pfizer Inc. (EP except GB, US, JP)
;  TITLE OF INVENTION: Neuropeptide receptor and uses thereof
;  FILE REFERENCE: PCS22032
;  CURRENT APPLICATION NUMBER: US/10/222,024
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: GB 0119920.7
;  PRIOR FILING DATE: 2001-08-15
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 361
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-222-024-2

  Query Match             19.0%;  Score 410.5;  DB 9;  Length 361;
  Best Local Similarity   30.2%;  Pred. No. 7.3e-26;
  Matches   99;  Conservative   58;  Mismatches  114;  Indels   57;  Gaps   10;

Qy        106 YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSI-CTTVFYTNLAIADFLFCVT 164
Db         24 YAHHSTARIVMPLHYSLVFIIGLVGNLLALVVIVQNRKKINSTTLYSTNLVISDILFTTA 83

Qy        165 LPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLP 224
Db         84 LPTRIAYYAMGFDWRIGDALCRITALVFYINTYAGVNFMTCLSIDRFIAVVHPLRYNKIK 143

Qy        225 KHTYALVTCGLVWATVFLYMLPFFI---LKQEYYLVQPDITTCHDVHNTCESSS-PFQLY 280
Db        144 RIEHAKGVCIFVWILVFAQTLPLLINPMSKQE-----AERITCMEYPNFEETKSLPWIL- 197

Qy        281 YFISLAFFGFLIPFVLIIYCYAAII-------------------RTLNAYDHRWLWYVK 320
Db        198 --LGACFIGYVLPLIIILICYSQICCKLFRTAKQNPLTEKSGVNKKALNTI--------- 246

Qy        321 ASLLILVIFTICFAPSNIILIIH------HANYYYNNTDGLYFIYL-IALCLGSLNSCLD 373
Db        247 --ILIIVVFVLCFTPYHVAIIQHMIKKLRFSNFLECSQRHSFQISLHFTVCLMNFNCCMD 304

Qy        374 PFLYFLMSKTRNHSTAYLTKXXNDLREQ 401
Db        305 PFIYFFACK------GYKRKVMRMLKRQ 326


RESULT 18
US-10-251-385-78
; Sequence 78, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
;  TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
;  TITLE OF INVENTION:  Protein-Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/10/251,385
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US/09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
;   LENGTH: 361
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-251-385-78

  Query Match             19.0%;  Score 410.5;  DB 9;  Length 361;
  Best Local Similarity   30.2%;  Pred. No. 7.3e-26;
  Matches   99;  Conservative   58;  Mismatches  114;  Indels   57;  Gaps   10;

Qy        106 YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSI-CTTVFYTNLAIADFLFCVT 164
Db         24 YAHHSTARIVMPLHYSLVFIIGLVGNLLALVVIVQNRKKINSTTLYSTNLVISDILFTTA 83

Qy        165 LPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLP 224
Db         84 LPTRIAYYAMGFDWRIGDALCRITALVFYINTYAGVNFMTCLSIDRFIAVVHPLRYNKIK 143

Qy        225 KHTYALVTCGLVWATVFLYMLPFFI---LKQEYYLVQPDITTCHDVHNTCESSS-PFQLY 280
Db        144 RIEHAKGVCIFVWILVFAQTLPLLINPMSKQE-----AERITCMEYPNFEETKSLPWIL- 197

Qy        281 YFISLAFFGFLIPFVLIIYCYAAII-------------------RTLNAYDHRWLWYVK 320
Db        198 --LGACFIGYVLPLIIILICYSQICCKLFRTAKQNPLTEKSGVNKKALNTI--------- 246

Qy        321 ASLLILVIFTICFAPSNIILIIH------HANYYYNNTDGLYFIYL-IALCLGSLNSCLD 373
Db        247 --ILIIVVFVLCFTPYHVAIIQHMIKKLRFSNFLECSQRHSFQISLHFTVCLMNFNCCMD 304

Qy        374 PFLYFLMSKTRNHSTAYLTKXXNDLREQ 401
Db        305 PFIYFFACK------GYKRKVMRMLKRQ 326


RESULT 19
US-10-251-385-206
; Sequence 206, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
;  TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
;  TITLE OF INVENTION:  Protein-Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/10/251,385
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US/09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 206
;   LENGTH: 361
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-251-385-206

  Query Match             19.0%;  Score 410.5;  DB 9;  Length 361;
  Best Local Similarity   32.2%;  Pred. No. 7.3e-26;
  Matches  102;  Conservative   58;  Mismatches  122;  Indels   35;  Gaps   11;

Qy        106 YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSI-CTTVFYTNLAIADFLFCVT 164
Db         24 YAHHSTARIVMPLHYSLVFIIGLVGNLLALVVIVQNRKKINSTTLYSTNLVISDILFTTA 83

Qy        165 LPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLP 224
Db         84 LPTRIAYYAMGFDWRIGDALCRITALVFYINTYAGVNFMTCLSIDRFIAVVHPLRYNKIK 143

Qy        225 KHTYALVTCGLVWATVFLYMLPFFI---LKQEYYLVQPDITTCHDVHNTCESSS-PFQLY 280
Db        144 RIEHAKGVCIFVWILVFAQTLPLLINPMSKQE-----AERITCMEYPNFEETKSLPWIL- 197
```

```
Qy        281 YFISLAFFGFLIPFVLIIYCYAAI----IRTL--NAYDHRWLWYVKAS---LLILVIFTI 331
Db        198 --LGACFIGYVLPLIIILICYSQICCKLFRTAKQNPLTEKSGVNKKAKNTIILIIVVFVL 255

Qy        332 CFAPSNIILIIH------HANYYYNNTDGLYFIYL-IALCLGSLNSCLDPFLYFLMSKTR 384
Db        256 CFTPYHVAIIQHMIKKLRFSNFLECSQRHSFQISLHFTVCLMNFNCCMDPFIYFFACK-- 313

Qy        385 NHSTAYLTKXXNDLREQ 401
Db        314 ----GYKRKVMRMLKRQ 326


RESULT 20
US-09-828-478-6
; Sequence 6, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
;  APPLICANT: Xiao, Yonghong
;  TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
;  TITLE OF INVENTION:  Protein
;  FILE REFERENCE: 04974.00458
;  CURRENT APPLICATION NUMBER: US/09/828,478
;  CURRENT FILING DATE:  2001-04-09
;  PRIOR APPLICATION NUMBER: 60/195,196
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: 60/254,876
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 367
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-828-478-6

  Query Match           19.0%;  Score 410.5;  DB 9;  Length 367;
  Best Local Similarity  30.7%;  Pred. No. 7.4e-26;
  Matches  107;  Conservative   48;  Mismatches  132;  Indels   61;  Gaps   12;

Qy         62 APPNSPEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKLIPAIYL 121
Db         35 APPGLITNFSLATAE----------QCGQE---------------TPLENMLFASFYL 67

Qy        122 LVFVVGVPANAVTLWMLFFRTRSICT--TVFYTNLAIADFLFCVTLPFKIAYHLNGNNWV 179
Db         68 LDFILALVGNTLALW-LFIRDHKSGTPANVFLMHLAVADLSCVLVLPTRLVYHFSGNHWP 126

Qy        180 FGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
Db        127 FGEIACRLTGFLFYLNMYASIYFLTCISADRFLAIVHPVKSLKLRRPLYAHLACAFLWVV 186

Qy        240 VFLYMLPFFILKQEYYLVQPDITTCHDVHNTCESSSPFQLY-------YFISLAFFGFLI 292
Db        187 VAVAMAPLLVSPQ-------TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 231

Qy        293 PFVLIIYCYAAIIRT----LNAYDHRWLWYVKASLLILVIPTICFAPSNI---ILIIHHA 345
Db        232 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 291

Qy        346 NYYYNNTDG--LYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
Db        292 SHGASCATQRILALANRITSCLTSLNGALDPIMYFFVAEKFRHALCNL 339


RESULT 21
US-10-092-135-7
; Sequence 7, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
;  TITLE OF INVENTION:  HGPRBMY27
;  FILE REFERENCE: D0134.NP
;  CURRENT APPLICATION NUMBER: US/10/092,135
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: US 60/273,808
;  PRIOR FILING DATE: 2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/278,983
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: RATTUS NORVEGICUS
US-10-092-135-7

  Query Match           19.0%;  Score 409.5;  DB 9;  Length 373;
  Best Local Similarity  28.9%;  Pred. No. 9.1e-26;
  Matches  111;  Conservative   64;  Mismatches  130;  Indels   79;  Gaps   13;

Qy         69 EFPPFSALEG----------WTGATITVKIKCPEESASHLHVKNATMGYLTSSLSTKL 115
Db          3 EVPWSAVPNGTDAAFLAGLGSLWGNSTI----------------ASTAAVSSSFRCAL 44

Qy        116 I---------PAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYTNLAIADFLFCVTL 165
Db         45 IKTGFQFYYLPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVLTL 104

Qy        166 PFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPK 225
Db        105 PALIFYYFNKTDWIFGDVMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLK 164

Qy        226 HTYALVTCGLVWATVFLYMLPFFILKQEYY----LVQPDITTCHDVHNTCESSSPFQLYY 281
Db        165 KKNAIYVSVLVWLIVVVAISPIL-----FYSGTGIRKNKTVTCYD-----STSDEYLRSY 214

Qy        282 FI-----SLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKASLLILVIF 329
Db        215 FIYSMCTTVAM--FCIPLVLILGCYGLIVRALIYKDLDNSPLRRKSIYLV---IIVLTVF 269

Qy        330 TICFAPSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
Db        270 AVSYIPFHVMKTMNLRARLDFQTPEMCDFNDRVYATYQVTRGLASLNSCVDPILYFLAGD 329

Qy        383 TRNHSTAYLTKXXNDLREQGQPSQ 406
Db        330 TFRRRLSRATRKASRRSEANLQSK 353


RESULT 22
US-09-827-937A-17
; Sequence 17, Application US/09827937A
; Patent No. US20020052043A1
; GENERAL INFORMATION:
;  APPLICANT: Li, Yi
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human G-Protein Coupled Receptors
;  FILE REFERENCE: 1488.1220003
;  CURRENT APPLICATION NUMBER: US/09/827,937A
;  CURRENT FILING DATE:  2001-04-09
;  PRIOR APPLICATION NUMBER: 08/852,824
;  PRIOR FILING DATE: 1997-05-07
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-827-937A-17

  Query Match           18.8%;  Score 405.5;  DB 10;  Length 348;
  Best Local Similarity  30.2%;  Pred. No. 1.8e-25;
  Matches   99;  Conservative   57;  Mismatches  115;  Indels   57;  Gaps   10;
```

```
Qy        106 YLTSSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSI-CTTVFYTNLAIADFLFCVT 164
Db         21 YAHHSTARIVMPLHYSLVFIIGLVGNLLALVVIVQNRKKINSTTLYSTNLVISDILFTTA 80

Qy        165 LPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLP 224
Db         81 LPTRIAYYAMGFDWRIGDALCRITALVFYINTYAGVNFMTCLSIDRFIAVVHPLRYNKIK 140

Qy        225 KHTYALVTCGLVWATVFLYMLPFFI---LKQEYYLVQPDITTCHDVHNTCESSS-PFQLY 280
Db        141 RIEHAKGVCIFVWILVFAQTLPLLINPMSKQE-----AERITCMEYPNFEETKSLPWIL- 194

Qy        281 YFISLAFFGFLIPFVLIIYCYAAII-------------------RTLNAYDHRWLWYVK 320
Db        195 --LGACFIGYVLPLIIIKICYSQICCKLFRTAKQNPLTEKSGVNKKALNTI--------- 243

Qy        321 ASLLILVIFTICFAPSNIILIIH------HANYYYNNTDGLYFIYL-IALCLGSLNSCLD 373
Db        244 --ILIIVVFVLCFTPYHVAIIQHMIKKLRFSNFLECSQRHSFQISLHFTVCLMNFNCCMD 301

Qy        374 PFLYFLMSKTRNHSTAYLTKXXNDLREQ 401
Db        302 PFIYFFACK------GYKRKVMRMLKRQ 323


RESULT 23
US-10-092-135-6
; Sequence 6, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
;  TITLE OF INVENTION:  HGPRBMY27
;  FILE REFERENCE: D0134.NP
;  CURRENT APPLICATION NUMBER: US/10/092,135
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: US 60/273,808
;  PRIOR FILING DATE: 2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/278,983
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-10-092-135-6

  Query Match             18.5%;  Score 399;  DB 9;  Length 373;
  Best Local Similarity   29.7%;  Pred. No. 6.6e-25;
  Matches  113;  Conservative   66;  Mismatches  143;  Indels   58;  Gaps   14;

Qy         52 PTLPIKT---FRGAPPNSFEEFPFSALEGWTGATITVKIKCPEESASHLHVKNATMGYLT 108
Db          7 PAVPNGTDAAFLAGPGSS----------WGNSTV----------ASTAAVSSSFKCALT 45

Qy        109 -SSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYTNLAIADFLFCVTLP 166
Db         46 KTGFQFYYLPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVLTLP 105

Qy        167 FKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTYRGLPKH 226
Db        106 ALIFYYFNKTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKK 165

Qy        227 TYALVTCGLVWATVFLYMLPFFILKQEYYLVQPDIT-TCHDVHNTCESSSPFQLYYFI-- 283
Db        166 KNAICISVLVWLIVVVAISP--ILFYSGTGVRKNKTITCYDT-----TSDEYLRSYFIYS 218

Qy        284 ---SLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKASLLILVIFTICF 333
Db        219 MCTTVAM--FCVPLVLILGCYGLIVRALIYKDLDNSPLRRKSIYLV---IIVLTVFAVSY 273

Qy        334 APSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNH 386
Db        274 IPFHVMKTMNLRARLDFQTPAMCAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRR 333

Qy        387 STAYLTKXXNDLREQGQPSQ 406
Db        334 RLSRATRKASRRSEANLQSK 353


RESULT 24
US-10-092-135-5
; Sequence 5, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
;  TITLE OF INVENTION:  HGPRBMY27
;  FILE REFERENCE: D0134.NP
;  CURRENT APPLICATION NUMBER: US/10/092,135
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: US 60/273,808
;  PRIOR FILING DATE: 2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/278,983
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: BOS TAURUS
US-10-092-135-5

  Query Match             18.5%;  Score 398;  DB 9;  Length 373;
  Best Local Similarity   28.5%;  Pred. No. 7.9e-25;
  Matches  111;  Conservative   66;  Mismatches  136;  Indels   76;  Gaps   15;

Qy         45 DTNNLAKPTLPIKTFRGAPPNSFEEFPFSALEGWTGATI--TVKIKCPEESASHLHVKNA 102
Db         14 DTAFLADP--------GSP--------------WGNSTVTSTAAVASPFKCA-------- 43

Qy        103 TMGYLT-SSLSTKLIPAIYLLVFVVGVPANAVTLWMLFFRTRSIC-TTVFYTNLAIADFL 160
Db         44 ----LTKTGFQFYYLPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFL 99

Qy        161 FCVTLPFKIAYHLNGNNWVFGEVLCRATTVIFYGNMYCSILLLACISINRYLAIVHPFTY 220
Db        100 YVLTLPALIFYYFNKTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKS 159

Qy        221 RGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY----LVQPDITTCHDVHNTCESSSP 276
Db        160 LGRLKKKNAVYISVLVWLIVVVGISPIL-----FYSGTGIRKNKTITCYDT-----TSDE 209

Qy        277 FQLYYFI-----SLAFFGFLIPFVLIIYCYAAIIRTLNAYD-------HRWLWYVKASLL 324
Db        210 YLRSYFIYSMCTTVAM--FCVPLVLILGCYGLIVRALIYKDLDNSPLRRKSIYLV---II 264

Qy        325 ILVIFTICFAPSNIILIIH-HANYYYNN------TDGLYFIYLIALCLGSLNSCLDPFLY 377
Db        265 VLTVFAVSYIPFHVMKTMNLRARLDFQTPEMCAFNDRVYATYQVTRGLASLNSCVDPILY 324

Qy        378 FLMSKTRNHSTAYLTKXXNDLREQGQPSQ 406
Db        325 FLAGDTFRRRLSRATRKASRRSEANLQSK 353


RESULT 25
US-10-024-494-30
; Sequence 30, Application US/10024494
; Publication No. US20030044898A1
;  GENERAL INFORMATION:
;       APPLICANT: LI, YI
;                  CAO, LIANG
;                  ROSEN, CRAIG A.
```